TEAM tAO

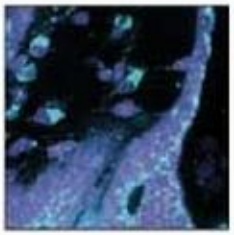
12 November 2004

Science

Vol. 306 No. 5699
Pages 1085–1240 $10

AAAS

Science



COVER Development of the light organ of the Hawaiian bobtail squid *Euprymna scolopes* is triggered by its luminescent bacterial partner *Vibrio fischeri* shortly after the symbiont colonizes epithelium-lined crypts deep within the organ. In this confocal micrograph, hemocytes can be seen infiltrating the host tissue that will undergo morphogenesis. See page 1186. [Image: T. Koropatnick]

Volume 306
12 November 2004
Number 5699

DEPARTMENTS

NEWS OF THE WEEK

NEWS FOCUS

1120

1143 & 1169

1144 & 1172

LETTERS

BOOKS *ET AL.*

POLICY FORUM

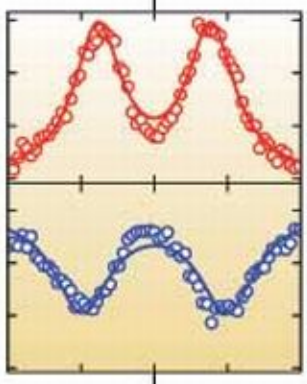
PERSPECTIVES

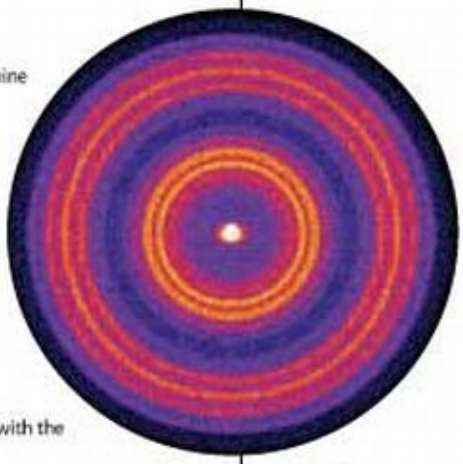
REVIEW

Contents continued

Science

SCIENCE EXPRESS www.sciencexpress.org

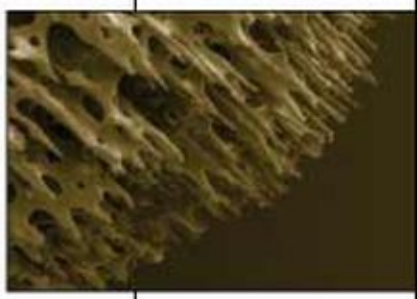




1158

BREVIA

REPORTS

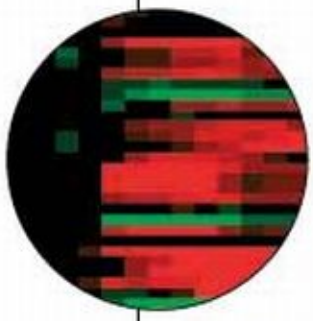


1161

Contents continued

Reports continued

1146 & 1180



1194

AAAS
ADVANCING SCIENCE. SERVING SOCIETY

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2004 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): $130 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $500. Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. Printed in the U.S.A.

Change of address: allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to Science, P.O. Box 1811, Danbury, CT 06813–1811. Single copy sales: $10.00 per issue prepaid includes surface postage; bulk rates on request. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $15.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075/05 $15.00. Science is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Contents continued

Science
www.scienceonline.org

sciencenow www.sciencenow.org Daily News Coverage

Sexual Selection Seen in Primate Genes
Gene that affects sperm coagulation evolved faster in promiscuous primates.

Global Warming Hits U.S. Ecosystems
First rigorous analysis shows climate change has many species on the move.

Target for Nicotine Addiction Found
Researchers gain new insight into the molecular mechanisms of addiction.

A career in solar materials.

science's next wave www.nextwave.org Career Resources for Young Scientists

Global/US: Academic Startup—SunPower *J. Kling*
Stanford's Richard Swanson has a solar materials business that is finally soaring.

Global/Netherlands: Priceless Materials—Diamond and Biosensors *T. Vrijenhoek*
Femke de Theije describes her Ph.D. experience and her path to a biosensor project at Philips.

UK: Starting Your Ph.D. *CareerDoctor*
Let the CareerDoctor guide you through the first steps toward earning your Ph.D.

Netherlands: Ph.D. Network—Universities Want Cheaper Doctorates *H. Obbink*
Universities may not be doing enough to lower costs for Ph.D. students. [in Dutch]

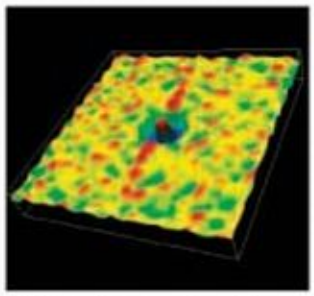
Career Development Center: Stuff and Bioethics *GrantDoctor*
The GrantDoctor writes about funding sources for materials science and bioethics research.

MiSciNet: Covering All the Bases *C. Parks*
Wayne Bowen, a minority biochemist and pharmacologist, discusses his research experiences.

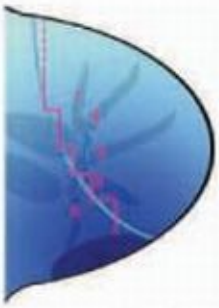
science's sage ke www.sageke.org Science of Aging Knowledge Environment

Perspective: Proteomics in Aging-Related Research *A. Gafni*
How many proteins does it take to age a cell?

News Focus: Out of Whack, Out of Mind *M. Leslie*
Misaligned brain cells might sap memory as we age.



Cell density map shows how memory stacks up.



A digit specification model.

science's stke www.stke.org Signal Transduction Knowledge Environment

Perspective: It Takes Time to Make a Pinky—Unexpected Insights into How SHH Patterns Vertebrate Digits *R. Zeller*
Digits are specified by intricate signaling interactions whereby cells sense the time of exposure to a morphogen and dynamically change their responsiveness to it.

ST on the Web
The Awards and Announcements section highlights 2004 Nobel Prize–winning cell signaling research.

Separate individual or institutional subscriptions to these products may be required for full-text access.

What's on the Label?

When you are buying food, are you one of the 30% of shoppers (an estimate in the United Kingdom) who always read the labels, or one of the 20% who rarely or never give them a glance? Do you know what to make of them if you read them?

Labels are meant to inform you and to help you to choose. But when you go shopping, how much time do you have to read about the differences between 30 types of chicken soup or 300 varieties of breakfast cereal? Consumers seem to want more and more choice, and consumer pressure groups definitely want more information on food labels. Choice and information are also attractive to regulators, because these options are less likely to be viewed as restricting individual freedom or stifling food industry innovation than the alternative of regulating food content.

In the United States, labeling regulations are largely about the material content. In Europe, the method and place of production may also be specified in law, even if they make no material difference to the contents. This difference in approach is evident in the labeling of genetically modified (GM) foods. Whether the plant from which a food is made is GM is irrelevant in the United States, given its emphasis on overall content rather than process. But in Europe, labeling of foods containing DNA or protein from GM plants is mandatory, and legislation has now been extended to include purified derivatives such as glucose syrup and canola oil (but not products from animals fed on GM animal feed or products made with GM technology, such as cheese).

Transatlantic differences in food labeling are also apparent when it comes to the biggest current challenge for food policy: obesity. Doing something about obesity is especially difficult for governments and regulators, because diet and lifestyle are in the territory of personal freedom, not state intervention. At the same time, the health care costs are potentially huge, so the pressure for action is on. The blend of action that is emerging, in both Europe and the United States, includes voluntary changes by the food industry, public education, and better labeling. Some countries and U.S. states are going even further; for instance, by restricting what can be sold in school vending machines and restricting television advertising. All of these changes are meant to make it easier for people to choose a healthy diet.

The world's fattest nation, the United States, has what is arguably the best nutrition labeling, with a mandatory nutrition facts panel. So would better labeling help? The largest food retailer in the United Kingdom, Tesco, has said that it plans to test a "traffic light" system, using red, yellow, and green colors to give consumers simple information about the main nutrients. Some object to this because of the potential implication that there are good (green) and bad (red) foods, whereas the traditional mantra from nutritionists is that there are only good and bad diets. But the food/diet distinction has changed as many people rely increasingly on ready-made meals or snacks. Research in the United Kingdom suggests that people would actually favor a simple sign-posting system such as traffic lights.

The food industry is responding to public interest in diet and health by making foods that claim to have specific health benefits. These come close to the border between food and medicine. You can buy cholesterol-reducing margarine, eggs that contain long-chain omega-3 unsaturated fatty acids, and yogurts that claim to help you balance your gut flora. The U.S. Food and Drug Administration has a three-tiered system for such health claims, depending on the strength of the evidence for the claim. The European Union does not have specific regulations, but plans to introduce rules within the next 2 years that will require the independent evaluation of health claims by the European Food Safety Authority. The implications of science-based regulation are enormous for the worldwide food industry, both because products that claim to improve your health are generally highly profitable and because, in the science of nutrition, there is often disagreement among experts. Over the next decade, increases in our understanding of the relationship between an individual's genetic makeup and his or her nutritional needs will open up a whole new area for debate about what goes on the label. The world of choice is not going to get any easier.

John Krebs

John Krebs is chairman of the Food Standards Agency, UK.

Published by AAAS

EDITORS' CHOICE

edited by Gilbert Chin

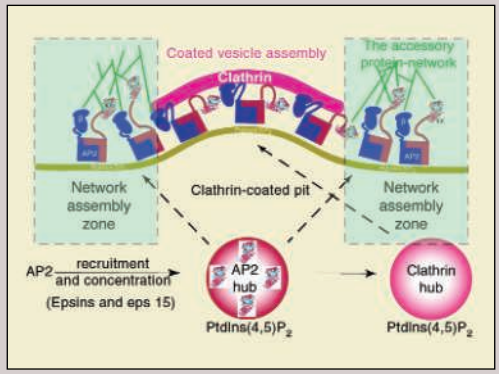
CELL BIOLOGY

Endocytosis at the Hub

In clathrin-mediated endocytosis, a network of proteins assembles on the cytoplasmic surface of the plasma membrane and promotes the pinching off of a membrane-bounded clathrin-coated vesicle. Together, the proteins select cargoes that are carried either inside the vesicle or in its membrane, modify the shape of the membrane, and drive invagination, vesicle scission, and eventual uncoating. A key player in this process is the AP2 clathrin adaptor protein, which is involved in concentrating selected cargo in the newly forming clathrin-coated pits.

In protein interaction networks, hubs are proteins that have disproportionately high numbers of interaction partners; in biological processes, hubs provide a temporal or spatial ordering to protein interactions. Praefcke *et al*. treat clathrin-mediated endocytosis as a module of a network and show how the α-appendage part of the AP2 protein works as an interaction hub. Only after being concentrated at sites of endocytosis do the appendages provide a multivalent binding platform (hub) for interaction partners (i.e., endocytic cargoes or other cargo adaptors). Thus, the partners will then be represented according to their relative affinities and concentrations in endocytic clathrin-coated pits and vesicles, even though any individual interactions will have been transient. — SMH

EMBO J. 10.1038/sj.emboj.7600445 (2004).

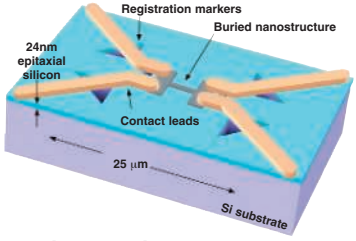


Interacting partners.

APPLIED PHYSICS

Registering Nanostructures

The manipulation of atoms using scanning tunneling microscopy (STM) has long promised the ability to fabricate nanometer- and atomic-scale electronic device structures. However, the realization of robust devices has been a difficult goal to attain simply because of the engineering problem of making electrical contact to the fabricated structure. The problem is that once the sample is removed from the ultrahigh vacuum where the STM atomic manipulation has taken place, the actual location of the structure is lost, particularly those regions that are buried under several layers of epitaxially grown semiconductor material. Ruess *et al*. have used a registration technique that allows the alignment of macroscopic electrodes to the nanoscale device elements buried underneath. The registration markers are etched into the substrate before the STM manipulation stage and so should be a general method for bottom-up fabrication of other nanoscale device structures. — ISO



Schematic showing contacts (yellow) to a buried 90-nm-wide quantum wire.

Nano Lett. **4**, 1969 (2004).

GEOLOGY

Taking Inventory

An enormous amount of methane, an important greenhouse gas, is stored in sediments in the ocean basins as icy methane clathrate and as gas trapped by this ice and by sediments. Catastrophic release of methane from this warehouse has been suggested to have caused abrupt climate change (warming) in the past, and there are concerns that a warmer future climate may destabilize this reservoir, which would enhance warming further.

To assess the amount of clathrate stored and to evaluate its stability, Buffett and Archer developed a mechanistic model for clathrate dynamics based on experimental and theoretical data on its stability and on factors affecting its formation and release, such as the supply of carbon to sediments and its diagenesis, storage, and oxygen content. Application to the current ocean basin implies that the global inventory is on the order of 10^{18} g of carbon stored as methane gas and clathrate. The modeling results imply that increasing temperature would likely deplete this inventory considerably; rebuilding would take several million years. The model also reveals that unless the oxygen content of the deep oceans was lower than at present, not enough methane would have been stored to account for the carbon isotope shift and the abrupt warming at the Paleocene-Eocene boundary, about 55 million years ago. — BH

Earth Planet. Sci. Lett. **227**, 185 (2004).

IMMUNOLOGY

Three in One

Vaccines are designed to generate robust immunity through the coactivation of the adaptive and innate arms of the immune system. This is achieved by steering tripartite responses to antigenic epitopes by helper T (T_H) cells, antigen-presenting dendritic cells (DCs), and antibody-producing B cells or the cytotoxic lymphocytes (CTLs) that ultimately execute pathogen clearance. However, the poor inherent immunogenicity of peptide epitopes favored in some vaccine formulas dictates the need for including complex and potentially toxic adjuvants that stimulate the essential priming activity of DCs.

Jackson *et al*. have refined this approach by synthesizing structures containing T_H epitopes coupled to B cell or CTL target epitopes. These were linked via a lipid moiety, which served to activate DCs through binding and activation of the innate signaling receptor TLR2. With different epitope

CONTINUED ON PAGE 1105

combinations, strong antibody and CTL responses could be elicited in models of viral and bacterial infection, as well as to tumors, and were comparable to responses to an adjuvant traditionally used in vaccines. The ability to combine adjuvant and antigenic properties in a single synthetic formula offers an attractive approach for future vaccine design. — SJS

Proc. Natl. Acad. Sci. U.S.A. **101**, 15440 (2004).

PHYSIOLOGY

Weight Control: It Takes a Village

About 250 million adults worldwide are obese, a condition that puts them at great risk for diabetes, heart disease, and other serious health problems. Although remarkable progress has been made in understanding the physiological and environmental factors that regulate body weight in mammals, much remains to be learned.

A new study in mice points to a surprising participant in body weight control: the community of bacteria (microbiota) that colonize the gut. Bäckhed *et al*. found that when they introduced the gut microbiota of normal mice into a special strain of "germ-free" mice, the recipients showed a 60% increase in total body fat within 2 weeks, even though they had eaten less and exhibited an increased metabolic rate. The microbiota appeared to promote fat storage by stimulating the synthesis of triglycerides in the liver and their deposition in adipocytes (fat cells). Based on their results, the authors hypothesize that changes in microbial ecology prompted by Western diets or differences in microbial ecology between individuals living in Western societies may affect predisposition toward obesity. — PAK

Proc. Natl. Acad. Sci. U.S.A. **101**, 15718 (2004).

CHEMISTRY

Keeping One's Head

Foam consists of a mass of bubbles—air trapped within thin liquid shells—that forms on agitation, such as when beer is poured into a glass. The stability of a foam depends on the nature of the liquid: evanescent in some cases, hardier in others because of additives that extend their lifetime. Surfactants, occasionally in conjunction with solid particles, are used as stabilizers because they reduce the surface tension of the liquid, preventing the bubbles from coalescing.

Alargova *et al*. have found that polymer microrods made from an epoxy-based photoresist can stabilize foams so that they resist collapse even when most of the liquid is allowed to evaporate. In contrast to foams made with the common household detergent sodium dodecyl sulfate, which survived for 2 days, the polymer rod foams were stable for more than 2 weeks. The authors speculate that the greater stability is due to two factors. First, the rods induce a much thicker liquid layer between the air bubbles, and these layers sterically repulse each other, thus preventing coalescence of the bubbles. Second, the rods within a layer form an intertwined network, thus increasing its overall strength and also imparting to the bubbles a spherical shape, which tends to be highly unstable in ordinary foams. — MSL

Langmuir 10.1021/la048647a (2004).

MICROBIOLOGY

Treasure Trove

Many bioactive small molecules were originally identified by screening extracts from microorganisms. These so-called natural products, some possessing medicinal value, then became the targets of structure determination and total synthesis. Traditional production methods depended on being able to pinpoint the source of the metabolite and to cultivate high-yielding strains of the isolated organism, but molecular biological advances have made it feasible to look directly for the genetic components of the biosynthetic pathways.

T. swinhoei.

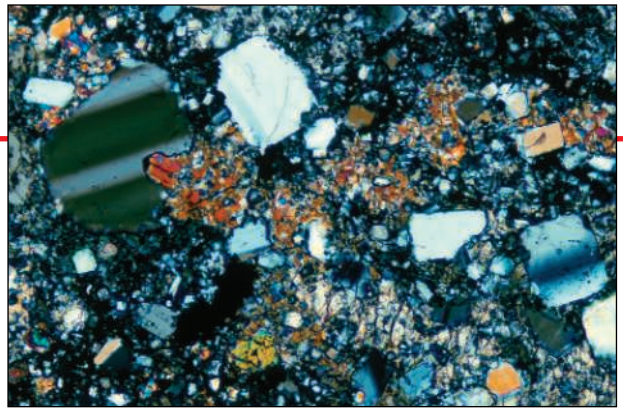
Piel *et al*. have extracted from the marine sponge *Theonella swinhoei* the gene clusters encoding the enzymes that make the polyketide onnamide A. Analysis of the gene structure indicates that their true source is probably an as-yet-unidentified bacterial symbiont, possibly of the *Pseudomonas* genus, harbored by the sponge. — GJC

Proc. Natl. Acad. Sci. U.S.A. **101**, 16222 (2004).

CREDITS: YOICHI NAKAO, UNIVERSITY OF TOKYO

NetWatch

edited by Mitch Leslie



IMAGES

Shooting the Moon

In the last year, the moon has put on a show for earthly observers, with two eclipses. If the events have whetted your appetite for lunar images, this pair of sites will allow you to explore the moon on large and small scales. The gallery* from the Lunar and Planetary Institute in Houston, Texas, supplies a digital version of NASA's classic 1971 atlas, a compilation of photos snapped by the Lunar Orbiter missions. Armchair astronauts can search for the 114-km-across H. G. Wells crater, the pockmarked Mare Australis, and other surface features. You can also browse the text of the original atlas.

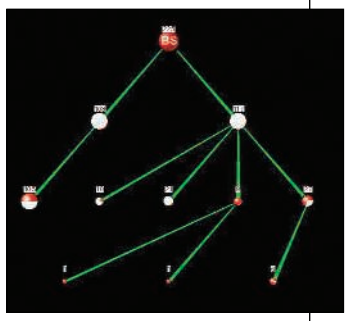
We think of the moon as gray, but under a microscope some of its rocks are surprisingly colorful. For a sample, check out this primer† on moon rocks and soil from geologist Kurt Hollocher of Union College in Schenectady, New York. The multicolored speckles above come from impact melt breccia, rock that partially melted when a meteorite or other wandering object slammed into the moon.

*www.lpi.usra.edu/research/lunar_orbiter

† www.union.edu/PUBLIC/GEODEPT/COURSES/petrology/moon_rocks/index.htm

TOOLS

Protein Sorter



Proteomic and genomic experiments pour out long lists of proteins. Researchers who need help comparing these proteins and figuring out what they do can open PANDORA, a protein-clustering tool hosted by the Hebrew University of Jerusalem in Israel. Users enter the proteins from their experiment, and then PANDORA gathers descriptions of the entries from other databases and uses them to parcel the molecules into smaller groups. The procedure "grabs the big picture," says co-creator Michal Linial of Hebrew University. For example, proteins that clump together in the analysis may also work together to perform a specific task or may congregate in the cell.

www.pandora.cs.huji.ac.il

EDUCATION

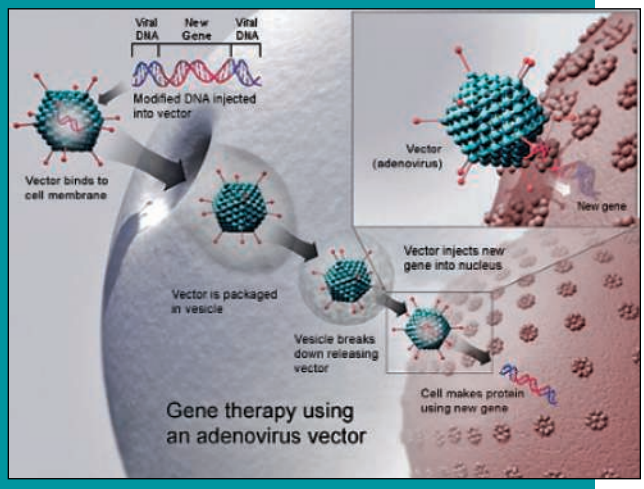
Chemistry Behind the Headlines

A researcher who submits a paper to a journal knows it has to pass the scrutiny of other scientists. The Web site Chemistry Is in the News gives students the chance to put their work through peer review while thinking and writing about science's role in current issues, from global warming to OxyContin addiction. Run by chemist Rainer Glaser of the University of Missouri, Columbia, and colleagues, the site provides guidelines to help students write reports about science-related stories that appear in the press. After exploring, say, the chemistry of the ozone-depleting pesticide methyl bromide and its possible effects on society, students can then post their efforts for evaluation by their classmates or students at other universities. In most science courses, says Glaser, "students are not challenged to think in broad terms and write about it." Teachers can apply to join the four universities already participating.

ciitn.missouri.edu/testsite/www/ciitn_main.html

EDUCATION

When Genes Go Bad



This primer on genetic diseases from the U.S. National Library of Medicine can serve as a reference for students and help teachers catch up on the latest findings. The goal of the Genetics Home Reference is to bridge a gap between researchers and genomics newbies, says project director Alexa McCray: "We were well aware of the wonderful things that have happened as a result of the human genome project, but there was no system that translated that information so that members of the public could understand it."

The handbook section explains topics such as inheritance, different kinds of mutations, genetic testing, and gene therapy. (Above, a virus toting modified DNA slips into a cell.) Users can learn about the genes responsible for illnesses and read up on some 100 conditions, from Alzheimer's disease (certain forms stem from mutations) to X-linked sideroblastic anemia, in which patients make too little hemoglobin. You can browse the descriptions by gene, condition, or chromosome. For readers who want to delve deeper, links lead to technical resources such as PubMed abstracts and gene reviews written for clinicians.

ghr.nlm.nih.gov/ghr/page/Home

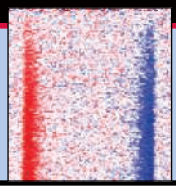
Send site suggestions to netwatch@aaas.org. Archive: www.sciencemag.org/netwatch

CREDITS (TOP TO BOTTOM): KURT HOLLOCHER/UNION COLLEGE; HEBREW UNIVERSITY OF JERUSALEM; NATIONAL LIBRARY OF MEDICINE

Published by AAAS

This Week

PAGE 1115
Setback for "open access"



1116
Spintronics comes into focus

U.S. SCIENCE POLICY

Bush Victory Leaves Scars—and Concerns About Funding

U.S. presidential science adviser John Marburger has some sharp words for researchers who publicly opposed President George W. Bush's reelection: Wrong message. Wrong audience. Wrong candidate.

Fresh from the election triumph by his boss and the Republican Party, Marburger warned last week in an interview with *Science* that criticism of the Administration's science policies during the campaign may be undermining public support for science. Offering a vigorous defense of the Administration's record, Marburger blamed critics for "looking at how the sausage is made" rather than at the product itself, which he characterized as a record windfall for science. Such partisan attacks, he suggested, may make it harder to prevent science from losing ground in the next 4 years given the demands of the war in Iraq, national security, and economic recovery.

Marburger's remarks came just 1 day after Bush described how he planned to "spend the political capital" from a 51% to 48% victory over Democrat Senator John Kerry and the increased Republican majority in both houses of Congress to reform Social Security, rewrite the tax code, and achieve other priorities. Science lobbyists are already worried about what will happen when Congress returns next week to finish the 2005 budget for the fiscal year that began 1 October. Their level of anxiety rises when they speculate about possible flat funding for key science agencies in the president's 2006 budget request this winter. And they may have to court new chairs of legislative panels that set policy and control budgets after a major reshuffling next year.

"Rightly or not, I think the science community is now perceived by this White House as the enemy, and that will make it harder to open doors," says physicist Michael Lubell, who handles government affairs for the American Physical Society. "It's one more factor in an increasingly complex situation," says David Moore of the Association of American Medical Colleges, who worries that fallout from the recent campaign could determine whether the Bush Administration "reaches out and engages [the science community] or goes in its own direction."

> "If we're not careful, the scientific community can become estranged from the rest of society."
>
> —John Marburger
> White House Science Adviser

If Marburger's analysis is correct, it's not the Administration but its scientific critics who have gone their own way, losing touch with society's concerns in the process. "Science needs patrons, and our patron is society," said the 63-year-old applied physicist, a former university president and head of Brookhaven National Laboratory. "But if we're not careful, the scientific community can become estranged from the rest of society and what it cares about."

Marburger said his remarks were directed at the 48 Nobel laureates who publicly endorsed Kerry last summer and a group—Scientists and Engineers for Change—that spent $100,000 to stage about 30 events on university campuses around the nation at which researchers criticized Bush's policies. The get-out-the-vote effort came on the heels of a fierce fight between the White House and some scientists, led by the Boston-based Union of Concerned Scientists (UCS), over allegations that the Administration has manipulated or suppressed science advice to advance its political agenda (*Science*, 9 April, p. 184). "I don't think that it was good for science to have done that," he says. "It was clear from the beginning because of the sweeping nature of the charges that the list of concerns were coming from the Democrats."

That's not true, says UCS chair Kurt Gottfried. "We're used to having our advice ignored or our recommendations rejected by both parties," says Gottfried, a physicist emeritus at Cornell University. "But we felt strongly that the quality of the scientific information coming from this Administration was being compromised by the way the process was being managed. And that's why we spoke up." Adds ecologist Jane Lubchenco of Oregon State University in Corvallis, "I can't speak for others, but I can assure you that my own motivation was not political."

Representative Sherwood Boehlert (R–NY), chair of the House Science Committee, thinks that everybody needs to take a deep breath. "Shame on both sides," says the self-proclaimed science booster, who this fall underwent triple bypass surgery but still managed to be reelected comfortably to a 12th term. "The rhetoric got a little bit excessive. I hope the Administration will demonstrate a greater degree of interest in the opinions of the scientific community. And I hope that scientists will realize that what they have been saying [about the Bush Administration] hasn't helped the profession."

Bush whacks. Supporters cheer the president's reelection over John Kerry.

For most rank-and-file scientists, the acid test for whichever party is in power is the flow of federal dollars into research. Not surprisingly, there is sharp disagreement between the president's supporters and his critics about how well science has fared in his first term even under that seemingly objective measure.

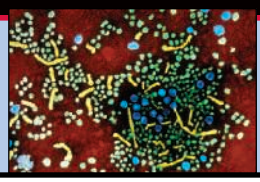
Throughout the cam- ▶

CREDITS (TOP TO BOTTOM): CHRIS MADDALONI; RON EDMONDS/AP PHOTO

Published by AAAS

1120
Music and evolution

1123
Flu models say vaccinate kids



1124
RNAi: not so specific after all?

paign and again last week, Marburger touted a 44% increase over 4 years in the government's overall research and development budget, from the $95 billion the Bush Administration inherited in 2001 to its 2005 request for $132 billion. Although defense-related research has led the way with a 62% hike, the National Institutes of Health and the National Science Foundation (NSF) have chalked up gains of 42% and 30%, respectively. "You really have to work at it to make a counterargument that science has not fared well in this Administration," says Marburger.

Critics see the numbers differently, however. In particular, they cite the government's failure even to begin the process of doubling NSF's budget, a promise written into a nonbinding 2001 law, and the fact that Congress has typically bulked up the Administration's initial request for NSF each year. They note that NIH's budget rose only 3% this year after a 5-year doubling that spanned the Clinton and Bush Administrations. And they say that most of the added defense spending goes to new weapons systems and fighting terrorism, not basic research.

Marburger and the Administration's critics may disagree over the record of the past 4 years, but both sides accept that the next few years will be tough for science budgets. Science's share of the overall discretionary spending budget (minus the mandatory payments and debt service that consume the ▶

California's Proposition 71 Launches Stem Cell Gold Rush

California is poised to leap ahead of the federal government as a backer of stem cell research after voters last week approved a 10-year, $3 billion plan to invest in the field. But the state is likely to proceed down a familiar path: Supporters of Proposition 71, which passed by a 59% to 41% margin, say the new California Institute for Regenerative Medicine will be modeled after the National Institutes of Health (NIH) in allocating its $295 million annual budget.

The state bond initiative, which will support work involving nuclear transfer (so-called research cloning), was backed by a staggering array of scientists and high-profile groups and received a last-minute endorsement from Republican Governor Arnold Schwarzenegger. Jubilant supporters are now moving on to the next phase, beginning with the selection within 40 days of a 29-member Independent Citizen's Oversight Committee. The governor and three top state officials will each appoint five members, and five University of California campuses will also have seats at the table. The likely chair is Donald Klein, the real estate magnate who led the campaign. In January the board will set up working groups on research funding, facilities, and standards, with the last being the first order of business, says campaign spokesperson Fiona Hutton. The first awards are to be made within 60 days of issuing interim standards.

Organizers promise that the new institute, at a site yet to be selected and with a permanent staff of 50, will be a first-class operation both ethically and scientifically. "The burden is upon us to prove that we are above reproach," says stem cell researcher Evan Snyder of the Burnham Institute in La Jolla. Stanford University stem cell researcher Irving Weissman, another leader in the campaign, expects universities to send either their presidents or medical school deans to represent them on the oversight committee, which he hopes will require reviewers from outside the state who are experts in the field.

California researchers will be able to put the voters' largess to good use, the initiative's backers assert. "Our goal has always been to mimic the NIH structure as much as possible," says Snyder. "We all come from the NIH tradition. ... We're not amateurs. We really know how science should be funded and conducted and administered."

Snyder says training will be a priority, and 10% of the budget will go to build research facilities and buy equipment. As for the rest, Weissman expects the bulk of the research funding to support investigator-initiated basic research rather than any top-down priorities set by the oversight committee.

Close to home. Hollywood producer Doug Wick and family celebrate passage of California's Proposition 71, which may one day help his daughter Tessa (far right) and others with diabetes.

One important early decision will be the selection of a full-time director. It "could easily be someone who's already an NIH administrator, with California roots," says Snyder. One person who fits that description is James Battey, who coordinates NIH's $214-million-a-year investment in human stem cell research, some $25 million of which goes to embryonic stem cells. Battey wouldn't comment, and an NIH spokesperson says NIH has nothing to say because Proposition 71 is "a state matter." But earlier this year Battey told *The New Yorker* that the measure "could have a really transforming effect on stem-cell research ... [and] would certainly make California an extremely attractive place to conduct it."

Despite the victory, supporters remain concerned about a bill sponsored by U.S. Senator Sam Brownback (R–KS) that would outlaw all forms of cloning. The new Senate includes four members whose records suggest they are likely supporters of the ban, which would leave Brownback only about five votes short of the 60 needed to move forward. At the same time, supporters of an opposing bill that would outlaw reproductive cloning but permit cloning for research remain hopeful that they, too, will prevail.

–CONSTANCE HOLDEN

CREDITS: FRAZER HARRISON/GETTY IMAGES

Published by AAAS

major part of the $2 trillion federal budget), at almost 14%, stands at a 40-year peak, Marburger notes, and sustaining such a "local maximum" is unlikely.

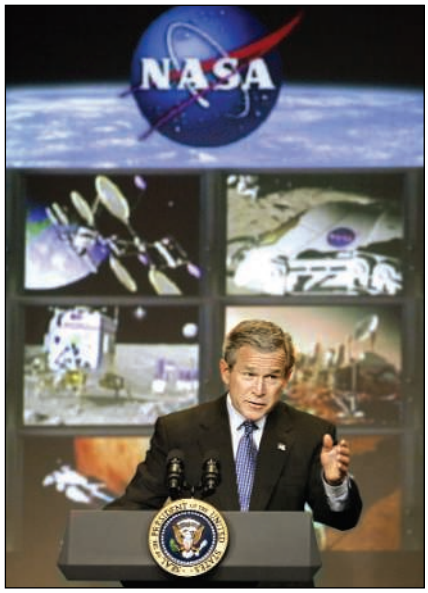
His co-chair of the President's Council of Advisors on Science and Technology, Silicon Valley financial guru Floyd Kvamme, goes one step further. He argues that science doesn't need a larger slice of the discretionary pie. "We're close to the limit of the amount of spending that the research enterprise can effectively absorb," says Kvamme, partner emeritus of the venture capital firm Kleiner, Perkins, Caufield, and Byers. "Of course, there are lots of other sectors demanding a share of that money. But I think we've reached a ceiling. It's unrealistic to talk about science taking up 15% or 20% of the discretionary budget."

The notion of a ceiling doesn't sit well with research advocates such as Nils Hasselmo, president of the 62-member Association of American Universities, science lobbyists, nor with Boehlert. "I think there's room for growth," says Boehlert, who helped shepherd the NSF reauthorization bill through Congress. Still, Boehlert says he expects "to be among those yelling" that the president's 2006 budget request for NSF and other science agencies, due out in February, is inadequate.

Next week Congress returns for a lame-duck session to complete work on the 2005 budget and, perhaps, conduct other business. But the real action won't begin until the more heavily Republican 109th Congress convenes in January. In the Senate, the GOP picked up four seats and holds a de facto margin of 55 to 45. (Senator James Jeffords of Vermont is an independent but usually sides with the Democrats.) In the 435-member House of Representatives, the GOP could have a 31-seat advantage after runoffs next month in Louisiana.

The new Senate is more likely to endorse the president's massive energy bill and overhaul the Endangered Species Act, but legislators may be more skeptical of his expansive—and expensive—moon/Mars exploration program, particularly at a time when private spaceships are capturing the headlines. The climate for science-related issues will also depend, in part, on who fills several key committee chairs. For example, Senator Ted Stevens (R–AK), who has shown little interest in regulating greenhouse gas emissions, is expected to become chair of the Commerce, Science, and Transportation Committee. He would replace Senator John McCain (R–AZ), who used the slot to push for climate change legislation and criticize the Administration's climate policies. Boehlert seems assured of remaining head of the equivalent House panel.

There will certainly be departures within the executive branch. The first science-related post to be vacant is at the Environmental Protection Agency, where research chief Paul Gilman announced that next month he will become head of a soon-to-be-announced research consortium of universities. NASA Administrator Sean O'Keefe is thought by many to be headed to another federal post, possibly even before this summer's planned shuttle launch, the first in Columbia's aftermath. And Health and Human Services Secretary Tommy Thompson, long rumored to be out the door after the election, said last week that "the president and I haven't had a chance to talk, and it's clearly up to the president." If he hits the road, one frontrunner for the job is Mark McClellan, a physician and economist who headed the Food and Drug Administration before moving to head Medicare. Lobbyists see him as more attuned to the interests of researchers than the former Wisconsin governor.



More space. Bush hopes Congress will fund his plans to explore the moon and Mars, announced earlier this year.

CREDIT: J. SCOTT APPLEWHITE/AP PHOTO

Despite the harsh language of the past year, Marburger says that "nothing has changed" in the way he will operate: "I'm going to continue to try to make sure that scientific issues are addressed." Harold Varmus, head of the Memorial Sloan-Kettering Cancer Center in New York City and one of the Nobelists who endorsed Kerry, expects to "still have conversations with people in government who hold opposing views." And NASA climate scientist James Hansen, who gave a widely publicized speech criticizing Bush just before Election Day, says "everybody is telling me to watch my back. … But I spoke from the heart, and these are honorable people."

–JEFFREY MERVIS

With reporting by Jocelyn Kaiser, Andrew Lawler, and David Malakoff

ScienceScope

Standards Set for Anthrax Detection Kits

Only one portable anthrax-detection kit of the five now on the market meets new standards established to help police and other first responders identify the deadly bacterium, an expert group says.

AOAC International (formerly the Association of Official Analytical Chemists) gave the word to the Department of Homeland Security (DHS), which paid for the analysis after DHS officials questioned the reliability of hand-held kits. AOAC's new standards require a kit to detect the presence of anthrax in a sample that contains at least 1 million anthrax spores and to distinguish accurately between anthrax and related organisms. According to AOAC, only the RAMP Anthrax Test, manufactured by Canada's Response Biomedical Corp., meets these criteria.

"This is going to have a major impact on the first-responder market" by improving reliability, predicts Calvin Chue, a pathogen-detection expert at the Johns Hopkins Center for Civilian Biodefense Strategies. Adds Stephen Morse, director of Columbia University's Center for Public Health Preparedness, "This is one nice step in the right direction."

–DAVID GRIMM

Wisconsin Academics Decry Move to Water Down Darwin

Wisconsin academics are rallying to reverse a decision last month by a local school board that would require students to "study various scientific models/theories of origins" rather than stick with Darwinian theory only.

The Grantsburg school board's action spurred Michael Zimmerman, dean of the College of Letters and Sciences at the University of Wisconsin, Oshkosh, to organize a flurry of letter writing by hundreds of scientists and theologians from universities around the state as well as high school science teachers. "We want to send as a strong a message as we can," says Zimmerman. Although Wisconsin state standards mandate the teaching of evolution, the board contends that the district has a right to make the standards more "inclusive."

Last month, the Dover Area School Board in Pennsylvania approved the teaching of "intelligent design" (*Science*, 5 November, p. 971). And a trial over an evolution "disclaimer" in textbooks is under way in Georgia. Says Eugenie Scott of the National Center for Science Education in Oakland, California: "After last Tuesday there are a lot of happy creationists around the country."

–CONSTANCE HOLDEN

Published by AAAS

U.S. HIGHER EDUCATION

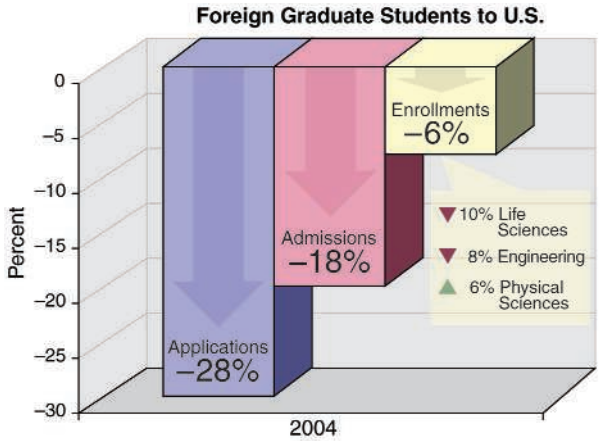
Decline in New Foreign Grad Students Slows

The number of international students beginning graduate studies at U.S. universities has declined for the third year in a row. But the 6% drop is the smallest in 3 years, an improvement that some attribute in part to faster handling of visa applications.

The news, in a survey released last week by the Council of Graduate Schools (CGS), comes as a relief to higher education organizations, which had braced for the worst earlier this year after a 28% drop in international graduate applications and an 18% drop in offers of admission. Enrollments, which represent the final step in that progression, were down 10% in the fall of 2003 following an 8% drop the year before. The decline appeared in the first academic year after the 11 September terrorist attacks and reversed several years of growth in the number of international students.

Universities have stepped up their efforts to assist foreign students, says CGS president Debra Stewart, "by streamlining their admissions processes, enhancing their use of technology, and forming international partnerships." The council says those measures contributed to a rise this year in the percentage of admitted international students who ended up enrolling.

"I am pretty sure that we have gotten over a hump in terms of visa delays," says Sherif Barsoum, associate director of the Office of International Education at Ohio State University in Columbus, who says he has received "not one e-mail or phone call this year complaining about a visa." Another survey released this week by five groups, including CGS and NAFSA: Association of International Educators, reported no change in undergraduate enrollments but declines in graduate enrollments at two-thirds of major research institutions.



Showing up. Enrollments are the last step in the process of attending graduate school, and trends vary by field.

"The good news is that the administration has become aware of the seriousness of the problem and has begun to take steps to address some of the obstacles that are discouraging or preventing legitimate students and scholars from coming to the United States," says NAFSA's Marlene M. Johnson. "The bad news is that, despite some positive signs, the overall numbers are still discouraging."

University administrators say their schools still need to combat the perception that U.S. campuses are unfriendly toward international students. Toward that goal, some universities are reimbursing students for the $100 fee the government charges to implement the Student and Exchange Visitor Information System, which tracks foreign students once they arrive. "It's not a lot of money, and it sends out a welcoming message," says Patricia Parker, assistant director of admissions at Iowa State University in Ames, which saw a 25% drop in first-time international graduate enrollment this fall.

Overall graduate enrollment is down 1%, according to the CGS survey, and 2% fewer domestic students are entering graduate school. The life sciences and engineering show the steepest declines in first-time international enrollment within the sciences, whereas the physical sciences are enjoying a 6% rise in first-time international students (see graph).

"We made offers to more international students this year than usual, anticipating that some of those who might accept would have trouble getting visas," says Allen Goldman, chair of the physics department at the University of Minnesota, Twin Cities. "That didn't materialize." The result, says Goldman, is a larger entering class—35 rather than 25 students—that is also more international.

–YUDHIJIT BHATTACHARJEE

OCEAN SCIENCE

NOAA to Retool Research Programs

The National Oceanic and Atmospheric Administration (NOAA) has embraced suggestions to consolidate some laboratories and make its funding practices easier to understand. But officials have rejected a proposal from an outside advisory panel for a new science czar.

"This looks like a very good start," said Science Advisory Board chair Leonard Pietrafesa, of North Carolina State University in Raleigh, reacting last week to NOAA's plan to shake up its $350 million research program, which includes everything from space-based climate and weather studies to fisheries research and deep-sea exploration. The outside panel, led by climate scientist Berrien Moore of the University of New Hampshire, Durham, suggested in August that the agency consolidate half a dozen laboratories in Boulder, Colorado, and revamp a convoluted external grants program. The panel also called for the agency to develop 5- and 20-year science plans and to put the agency's research programs under the control of a new senior administrator and an allied advisory board (*Science*, 11 June, p. 1579).

On 3 November, NOAA officials told the board that they have nixed the last two ideas because of congressional opposition to any more bureaucratic layers. Instead, NOAA Deputy Administrator James Mahoney says his position has been "restructured" to increase his oversight of research. Officials said they expect other changes will take place over the next 18 months, including clarifying both the amount of money available and the application process for extramural researchers and creating a Web-based grants management system. "I don't think anyone would give NOAA an 'A' for our involvement with outside" researchers, Mahoney said. He also promised to increase the number of administrators overseeing key science programs, saying that although NOAA has "an abundance of capable researchers," the administrative corps "is very thin indeed."

Perhaps the biggest question mark is the fate of NOAA's six Colorado laboratories. Congressional critics have argued for lumping the labs together into fewer and less expensive units, but researchers worry that the move could hurt science programs. Mahoney says a task force could issue a consolidation plan as early as this fall, followed by another group looking at NOAA's ecological research programs. Any proposed changes, however, must survive vetting from the White House Office of Management and Budget and win the support of Congress.

–DAVID MALAKOFF

SOURCE: CGS

Published by AAAS

SCIENTIFIC PUBLISHING

Mixed Week for Open Access in the U.K.

CAMBRIDGE, U.K.—Supporters of "open access" scientific publishing—in which authors pay the cost of publication and accepted papers are freely available online—have received a public setback and a private boost in the United Kingdom in the past few days.

The British government, saying it is "not obvious … that the 'author pays' business model would give better value for money than the current one," rejected recommendations from the House of Commons Science and Technology Committee to fund some costs associated with open-access publishing. The committee promptly accused the government of buckling under pressure from scientific publishers. On the other hand, the Wellcome Trust—the largest funder of basic biomedical research in the United Kingdom—threw its considerable weight behind open-access publishing. It announced that it will require researchers it funds to deposit papers in a public archive "within 6 months of publication," and it is discussing the creation of a European version of PubMed Central, the open archive of published papers run by the U.S. National Library of Medicine.

Open up. Parliamentary committee chair Ian Gibson (*left*) and Wellcome Trust chief Mark Walport back public access.

CREDIT (RIGHT): THE WELLCOME TRUST

The government's statements came in a detailed response to a report issued this summer by the House of Commons committee (*Science*, 23 July, p. 458). The committee concluded that open-access publishing is an "attractive" idea but called for further study because of the possible impact on learned societies, which rely financially on journal subscriptions. The panel also was concerned that the pharmaceutical industry, which subscribes to many journals but contributes few papers, would get a free ride. In the meantime, the committee recommended that the research councils create a fund to which authors could apply for the costs of publishing in open-access journals. It also called for U.K. universities to set up online repositories for their researchers' preprints and published papers, which would be posted at some agreed point after publication in a journal, and called for the government to set up a body to coordinate these archives.

The true costs of open-access publishing are still not clear, the government responded, noting that "before fully supporting any new business model, the Government will need to be convinced that this model is better and cheaper." It declined for now to require that government-funded researchers deposit their published papers in open repositories or to establish a fund to pay authors' publication fees.

The House of Commons committee published the government's response on 8 November, along with its own commentary* accusing the Department of Trade and Industry (DTI) of neutralizing dissenting voices within the government. "DTI is apparently more interested in kowtowing to the powerful publishing lobby than it is in looking after the best interests of British science," says committee chair Ian Gibson. The government's response is "a defence of the status quo," adds Jan Velterop, publisher of the author-pays journal *Biomed Central*.

While the government was urging caution on open-access publishing, the Wellcome Trust was stepping up its support. The trust will now fund "reasonable" costs for its grantees to publish in author-pays online journals; in addition, it will require grantees to deposit published papers in a public archive. Trust director Mark Walport estimates that publication charges will only amount to 1% of the trust's research costs. The goal is "to achieve maximum value from our research through maximum distribution," says Walport.

Gibson says the committee will continue the fight. Librarians are "gung-ho" about public access, he says, and he hopes that the research councils will soon come out in favor too. In the new year, Gibson says, there will be a debate in the House of Commons: "We're going to argue this with them."

–DANIEL CLERY

*www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/1200/1200.pdf

ScienceScope

High Demand Leads to Shortage of Malaria Drug

The World Health Organization (WHO) cautioned last week that supplies of a potent antimalaria drug may fall up to 4.5 million doses short of their demand forecasts. Officials at Novartis, the Swiss-based company that manufactures the drug, blamed the shortfall on agricultural suppliers failing to keep up with growing demand in developing countries.

Artemether-lumefantrine (brand name Coartem) is a form of artemisinin-based combination therapy (ACT) favored by WHO because of the increasing number of strains of malaria that are resistant to traditional drugs (such as chloroquine). Since 2001, Novartis has provided Coartem to developing countries at cost.

A key ingredient of the drug comes from the Chinese wormwood plant (*Artemisia annua*). Suppliers of the ingredient have struggled to ramp up production to meet the growing interest in ACT, says Andrew Bosman, a medical officer at WHO. The plant takes 6 months to cultivate, and the drug requires 3 to 5 months to process, resulting in a mismatch between orders and supplies, he says. To help combat the shortage, WHO will step up its malaria-prevention efforts and plans to develop a prioritization system for drug distribution.

–SEAN BRUICH

IBM Study Challenges Cancer Claims

An IBM-funded study has found that the company's workers face no greater risk than the general population of developing cancer. The results, which have not yet been peer reviewed, contradict an earlier study paid for by attorneys of former IBM workers suing the company for causing their cancers (*Science*, 14 May, p. 937).

Researchers at Harvard University's School of Public Health and the University of Alabama examined health records of 126,000 workers and compared overall cancer death risks between IBM workers and the general population. The workers had a 16% lower risk of cancer and a 35% lower risk of dying than the general population, according to a memo released last week.

Joseph LaDou, director of the University of California, San Francisco's International Center for Occupational Medicine, who gave unpaid advice to lawyers for former IBM workers suing the company, believes IBM biased the study and says the results could reflect the fact that manufacturing workers tend to be healthier than the population at large. IBM representatives did not return repeated calls seeking comment.

–DAN FERBER

Published by AAAS

PALEOANTHROPOLOGY

Skeptics Question Whether Flores Hominid Is a New Species

When a research team announced last month that it had found a new species of 18,000-year-old tiny human in a cave on the Indonesian island of Flores, it seemed almost too amazing to be true (*Science*, 29 October, p. 789). Now a small but vocal group of scientists argues that the skeleton dubbed *Homo floresiensis* is actually a modern human afflicted with microcephaly, a deformity characterized by a very small brain and head. Meanwhile, an Indonesian scientist who also challenges the skeleton's status has removed the skull to his own lab for study. But members of the original team of Australian and Indonesian scientists staunchly defend their analysis, and outside experts familiar with the discovery are unmoved by the critique.

The main challenge comes from paleopathologist Maciej Henneberg of the University of Adelaide in Australia and anthropologist Alan Thorne of the Australian National University in Canberra. Neither has seen the specimen itself, and as *Science* went to press, they had yet to publish their criticisms in a peer-reviewed journal. But Henneberg published a letter in the 31 October Adelaide *Sunday Mail* arguing that the skull of the Flores hominid is very similar to a 4000-year-old microcephalic modern human skull found on the island of Crete. And at a press conference on 5 November, Indonesian paleoanthropologist Teuku Jacob of Gadjah Mada University in Jakarta claimed that the specimen was a diminutive modern human. Jacob, once described as the "king of paleoanthropology" in Indonesia (*Science*, 6 March 1998, p. 1482), has had the skull transported to his own lab from its original depository at the Center for Archaeology in Jakarta, according to center archaeologist Radien Soejono, who is a member of the original discovery team.

In its original paper, the team considered and rejected several possible deformities, including a condition called primordial microcephalic dwarfism (*Nature*, 28 October, p. 1055). But Henneberg claims that the authors failed to consider a related condition called secondary microcephaly. "They jumped the gun," he told *Science*. Henneberg, who with Thorne favors a multiregional model of human origins that some say is at odds with the finding of a distinct but recent human species on Flores, concludes that the skeleton is "a simple *Homo sapiens* with a pathological growth condition." (Multiregionalism holds that modern humans evolved after 2 million years of interbreeding among worldwide populations; the evolution of a distinct species would require a long period without interbreeding).

But archaeologist Michael Morwood of the University of New England in Armidale, Australia, a leader of the team that discovered the skeleton, insists that the skeleton is not a pathological case. "We now have the remains of at least seven individuals," he says. "All are tiny, and all can be referred to as *Homo floresiensis*."

The team is backed by several outside researchers. Anthropologist Leslie Aiello of University College London says the skeleton cannot be that of a modern human because the postcranial bones indicate a separate species. "The pelvis is virtually identical to that of an australopithecine," much wider than the modern human pelvis, she says. And compared with modern humans, "the arms are long in relation to the legs." Chris Stringer of the Natural History Museum in London sums up many researchers' opinions by saying, "This cannot be a peculiar modern human."

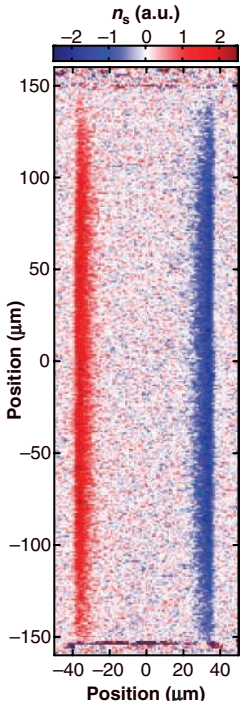
Surprising skull. A few researchers say the Flores skull may be a deformed modern human.

–MICHAEL BALTER

CONDENSED MATTER PHYSICS

Spin Current Sighting Ends 35-Year Hunt

The electron's charge gets all the glory: It is, after all, responsible for the plethora of electronic gizmos that surround us. But the particle's magnetic behavior—a property known as spin—has also been tantalizing scientists for decades. Thirty-five years ago, for example, Russian theorists suggested that impurity atoms in a semiconductor might interact with electrons' spins to redirect currents flowing through it. A related effect, called the Hall effect—in which magnetic fields push electrons around by interacting with their charge—had been known for more than a century. But despite decades of work, the spin-based Hall effect had never been spotted—until now.



Sidetracked. Impurity atoms deflect electrons with different spins (red and blue) to opposite sides of a semiconductor chip.

In a paper published online this week by *Science* (www.sciencemag.org/cgi/content/abstract/1105514), researchers led by David Awschalom, a physicist at the University of California, Santa Barbara (UCSB), report the first experimental sighting of the spin Hall effect. "It is as beautiful as it is a breakthrough experiment," says Gerrit Bauer, a theoretical physicist at the Delft University of Technology in the Netherlands. Daniel Loss, a theoretical physicist at the University of Basel, Switzerland, agrees. "The data is very clear," he says. "It's very impressive." The new scheme works in standard semiconductors widely used in industry today. That could be a major boon to the nascent field of spintronics, which promises to create a new class of high-speed, low-power electronic devices that manipulate the spin of electrons.

An American physicist named Edwin Hall discovered the original Hall effect in 1879. The effect occurs when an electric current moves through a metal strip while a magnetic field is applied top down through the metal. The magnetic field interacts with the charge of the moving electrons, deflecting some to the left and some to the right sides of the strip. In 1971, Mikhail Dyakonov and Vladimir Perel of the Ioffe Physico-Technical Institute in Leningrad suggested that electrons' spins might trigger similar detours. These spins behave like ▶

CREDITS (TOP TO BOTTOM): P. BROWN/UNIVERSITY OF NEW ENGLAND, ARMIDALE, AUSTRALIA; Y. K. KATO *ET AL.*

tiny bar magnets that point up or down. Dyakonov and Perel proposed that electrical defects in semiconductors could create a localized electromagnetic field that would shunt spin-up and spin-down electrons to opposite sides of the semiconductor, a scheme that came to be known as the "extrinsic" spin Hall effect. Decades later, other theorists suggested that such deflection might also result from an "intrinsic" effect due to the strain between atoms in a semiconductor alloy. In recent years, theorists have clashed sharply over which effect would more likely be spotted.

Awschalom says his team waded into this battle somewhat by accident. Earlier this year, the Santa Barbara researchers, who included graduate students Yuichiro Kato and Roberto Myers along with electrical engineer Art Gossard, discovered a scheme for electrically injecting spins into a semiconductor, another long-sought goal of spintronics (*Science*, 2 April, p. 42). They were tracking spins by a technique called scanning Kerr microscopy, in which researchers bounce polarized laser light off a semiconductor sample. If electrons on the surface atoms all have spins in one preferred orientation, the polarization of the ricocheting photons will rotate slightly.

When the UCSB researchers used a Kerr microscope lens with 1-square-micrometer resolution, they saw clear bands of electrons with opposite spins huddled along the sides of the semiconductor chips. All it took to create the bands was to push an electric current through the semiconductor chip.

The researchers first detected the bands in a semiconductor chip made from gallium arsenide (GaAs), similar to the chips in cell phones. Then, in hopes of resolving the battle over intrinsic and extrinsic effects, they looked for the effect in a semiconductor called indium gallium arsenide (InGaAs). Theorists had suggested that the atomic bonds in InGaAs had the right sort of strain to produce a sizable intrinsic spin Hall effect even if the impurities didn't play a role. But the researchers found that impurities could explain virtually all of the buildup of spins they saw. "It looks like [the extrinsic effect] is the major player" in the semiconductors studied so far, Loss says.

Using standard techniques for tailoring the amount of impurity atoms and other "defects" in semiconductors, "it should be possible to engineer materials to increase the size of this effect," Awschalom says. That in turn could point the way for spintronics researchers to develop an array of spin-manipulating devices to switch currents of particular spins on and off, as well as steer, filter, and amplify them. That might be enough to finally bring the electron's spin a little limelight of its own. **–ROBERT F. SERVICE**

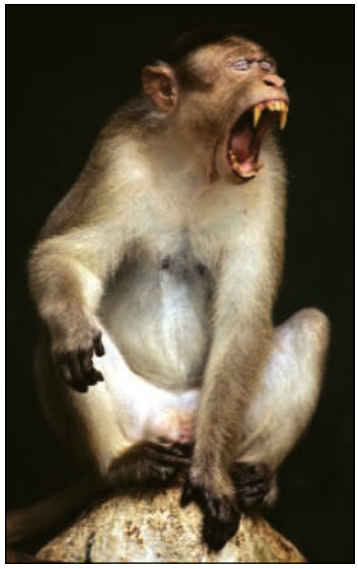
CREDIT: N. ROWE/THE PICTORIAL GUIDE TO THE LIVING PRIMATES

MEDICINE

Sperm-Targeting Vaccine Blocks Male Fertility in Monkeys

When the 1880s debut of vulcanized rubber and the 1930s advent of latex mark the latest advances, one quickly understands the sorry state of male contraceptives. Researchers in this beleaguered field have tried to supply men with other options but have had little success. Now, on page 1189 of this issue of *Science*, a team in the United States and India reports preliminary results of its new contraceptive vaccine for males. Although not without problems, the vaccine prevented pregnancy in female partners of the male monkeys receiving it.

"There seems to be some real promise," says Ronald Swerdloff, a reproductive endocrinologist at the University of California, Los Angeles. Still, "it's just early in the game," with too few monkeys tested, to conclude whether the approach will pan out, he adds.



Sperm stopper. Male contraceptive vaccine works in monkeys.

Reproductive biologist Michael O'Rand of the University of North Carolina, Chapel Hill, crafted the vaccine several years after reporting his discovery of a novel male-only protein in 2001. The protein, called Eppin, has been found so far on the surface of sperm cells and elsewhere in the testis and the epididymis. Its function isn't clear. But drawing on the general strategy of immunocontraception, in which vaccines are designed to act as contraceptives, O'Rand reasoned that if a male harbored antibodies to this protein, his sperm might malfunction.

O'Rand teamed up with colleagues at the Indian Institute of Science in Bangalore, which hosts a large primate research center. There, the group vaccinated six monkeys with human recombinant Eppin protein and administered a sham vaccine to six others. The team hypothesized that the monkeys needed high levels of antibodies to Eppin in their blood for the vaccine to work, especially because antibody levels drop in the epididymal fluid. So when only four of the six treated monkeys displayed antibody levels that O'Rand's team deemed sufficient, the other two animals were dropped from the study. The team brought in three additional monkeys, who also reached the desired antibody levels. It's not clear why some monkeys did not produce sufficient antibodies, a problem other immunocontraceptives have encountered.

The monkeys received boosters of vaccine every 3 weeks. Because the vaccine didn't lower sperm count or alter sperm in an easily detectable way, scientists resorted to another method of testing its effectiveness: The immunized male monkeys spent several days each with three different females during the fertile peak of the females' menstrual cycle. The upshot: None of the seven vaccinated monkeys managed to impregnate a female. Four of the six control monkeys did.

The contraceptive effect of the vaccine was intended to be reversible; once the booster shots were stopped, the researchers anticipated that antibodies to Eppin would decline and fertility would return. But only five of the seven vaccinated monkeys, some of whom received booster shots for nearly 2 years, recovered their fertility during the course of the study. "It's hard to say" what that means, says O'Rand. "Maybe they recovered 2 weeks after we quit" testing them. Although conceding that the vaccine has a long way to go, O'Rand believes the study offers "a proof of principle" for immunocontraception, which has been so relegated to the sidelines that the National Institutes of Health no longer funds research on it.

Companies have also been hesitant. New Jersey–based Organon studied immunocontraception for females before backing out, says Willem de Laat, the company's medical director. Instead, Organon and another company, Schering in Berlin, are testing a combination of oral progestin and injected testosterone.

O'Rand and his colleagues, heartened by what they consider a success, are now trying to understand how, exactly, their vaccine disrupts fertility. One possibility is that the technique leaves sperm sluggish. Whatever its mechanism, O'Rand and other contraceptive researchers hope the new vaccine will provide a shot in the arm for the field.

–JENNIFER COUZIN

DRUG TESTING

Regulators Talk Up Plans for Drug Biomarkers ...

New methods of predicting clinical outcomes are getting serious consideration at the U.S. Food and Drug Administration (FDA). Last week the agency invited members of one of its advisory committees to a meeting in Washington, D.C., where attendees explored the use of biomarkers to monitor everything from protein levels to bone density, to gauge a drug's effectiveness, or possibly to measure the progression of a disease. FDA officials at the meeting said they're launching a multiyear effort to bring biomarkers into the mainstream of drug discovery.

Architect. FDA's Lawrence Lesko is pushing biomarkers forward.

"There are some significant payoffs if this is successful," such as speedier drug trials, says Lawrence Lesko, FDA's director of clinical pharmacology and biopharmaceutics. Still, Lesko has encountered hesitancy: "People are sensitive to past failures" of biomarker use, he says.

The new effort began with former FDA commissioner Mark McClellan, who left the agency in February and has not yet been replaced. McClellan made biomarkers part of the agency's "Critical Path" initiative, a plan released in March to speed drugs to market.

Currently, most drugs are approved based on so-called clinical endpoints, such as longer survival for cancer drugs or fewer fractures for an osteoporosis drug. Researchers have long believed that there are reliable surrogates that can be detected earlier for many clinical endpoints. For example, the time it takes a cancerous tumor to resume growing during or after a specific type of treatment may indicate how long a patient will survive. Biomarkers tied to a clinical outcome, like this one, could be used as surrogate endpoints in trials.

FDA already approves some drugs based on surrogates, particularly for life-threatening diseases like AIDS (for which it has used surrogates since 1992) and cancer. But the history of biomarkers is marred by some high-profile disasters. One was the widespread use of two antiarrhythmic drugs, encainide and flecainide. These drugs were intended to reduce the likelihood of a second heart attack because uncontrolled arrhythmia was considered a predictor. When a clinical trial actually tested the drugs for this indication in the late 1980s, three times as many people died in the drug arm as in the placebo group, and the study was halted.

Advocates nevertheless argue that using biomarkers will do more good than harm by getting new medicines to patients quickly. "We're putting these molecules in these painfully slow, archaic programs" for testing, says Paul Watkins, a liver specialist at the University of North Carolina, Chapel Hill.

Although FDA is enthusiastic, it remains vague about how it might change its methods. As a first step, says Lesko, it will set up an internal working group to discuss what qualifies a biomarker as a surrogate endpoint. It may also comb through archival data for promising biomarkers. "We have to get down to some more specifics … to make this proposal come to life," Lesko admits. Working closely with FDA's acting commissioner for operations Janet Woodcock, he says he hopes to foster collaborations between industry, academia, and FDA to get ideas moving beyond the basic-research stage.

But pushing biomarkers forward is not risk-free. "As we study biomarkers, we're going to develop evidence that impugns their use," said John Wagner, senior director of clinical pharmacology at Merck Research Labs, at last week's meeting. And Watkins, who heads a new government-funded consortium on drug-induced liver toxicity, notes that biomarkers will gain acceptance only if they're used to flag issues of safety as well as efficacy. In May, the consortium began enrolling the first of dozens of patients who suffered severe liver effects from one of four drugs, along with matching controls. It aims to link molecular markers with susceptibility to this common side effect.

"Everybody wants to sit down and use something that's predictive" of clinical outcome, says Charles Grudzinskas, a former drug industry executive and founder of NDA Partners LLC, a Washington, D.C., consulting firm. He adds that FDA seems ready to lead the way with an attitude that, "we're willing to go out on skinny branches." But he and others are waiting to see whether FDA's next chief will throw the agency's prestige and funds behind this cause. **–JENNIFER COUZIN**

CREDIT: FDA

CANCER RESEARCH

... And NCI Hears a Pitch for Biomarker Studies

Cancer researcher Lee Hartwell this week proposed a major new initiative to discover biomarkers for the early detection of cancer. In a white paper presented to the National Cancer Institute's (NCI's) Board of Scientific Advisors, Hartwell, of the Fred Hutchinson Cancer Research Center in Seattle, Washington, outlined how the project would scan thousands of blood samples from cancer patients for proteins and other biological molecules that can indicate incipient tumors.

Hartwell said his plan would boost spending on biomarkers, now overshadowed in NCI's budget by new drug development and even prevention trials. Although only a handful of biomarkers are widely used, the sequencing of the human genome and the debut of new, automated mass spectrometry machines for detecting proteins leaves the field ripe for new breakthroughs, Hartwell said: "We need to organize our efforts." His proposed "Coordinated Clinical Proteomics and Biomarkers Discovery Initiative" would include centers for testing biomarker technologies, a repository of reagents, and a public proteomics software package.

The initiative, for which Hartwell offered no price tag, would complement a technology plan outlined by the Broad Institute's Eric Lander in September (*Science*, 24 September, p. 1885). Both proposals were requested by NCI Director Andrew von Eschenbach as part of his plan to eliminate death and suffering from cancer by 2015.

Board members peppered Hartwell with questions. Because the same protein can occur in many forms, for example, "you have to realize it's going to be much more complicated than looking for a single protein," said Susan Horwitz of Albert Einstein School of Medicine in New York City. Richard Schilsky of the University of Chicago suggested that costs might approach those of drug testing if each new marker had to be validated clinically. Hartwell disagreed, saying the project would "piggyback" on other large studies such as the Women's Health Initiative by borrowing tissue samples. "I don't see the validation adding a great deal of expense," he said.

Several members also asked how the initiative would fit with existing NCI programs and the National Institutes of Health's broader Roadmap, which includes proteomics. Von Eschenbach responded that it would "dovetail" with them but did not specify how. Questions about one such detail—how NCI would pay for a new biomarker plan with an ever-tightening budget—may come up later this month at a meeting of the National Cancer Advisory Board. **–JOCELYN KAISER**

Published by AAAS

Why did the ability to carry a tune evolve? At an unusual, high-level meeting, researchers pondered whether music helped our ancestors survive and reproduce or whether it is merely a happy evolutionary accident

Seeking the Key to Music

READING, ENGLAND—On a recent fall evening, the lobby of the archaeology building at the University of Reading was the scene of a strange ritual. Twenty-five researchers danced in a circle while blowing on the ends of differing lengths of rubber tubing. Pedro Espi-Sanchis, a music educator based in South Africa, had cut the tubing such that the notes produced by the pieces spanned two full octaves. Espi-Sanchis encouraged everyone to toot to his or her own inspiration, but to try not to repeat what others were doing. After several minutes, to everyone's delighted surprise, the individual notes coalesced into a single pleasing melody to which the dancers swayed and dipped in rhythm.

This spontaneous musical performance, a highlight of a recent workshop on the evolution of music and language,* illustrated one of the meeting's key themes: Music, like language, can be a form of communication and coordination among people. Moreover, music is an exquisitely powerful way of conveying emotion, a task at which language all too often falls short.

Yet although few researchers question that human language arose by means of natural selection, presumably because more accurate communication helped early humans survive and reproduce, the evolutionary significance of music has remained open to debate. The meeting, organized by Reading archaeologist Steven Mithen and music educator Nicholas Bannan, was intended as a first step in setting a research agenda to explore the evolution of music.

In 1997, cognitive scientist Steven Pinker, then of the Massachusetts Institute of Technology, threw down the gauntlet in his book *How the Mind Works*, when he suggested that music itself played no adaptive role in human evolution. Rather, Pinker argued, music was "auditory cheesecake," a byproduct of natural selection that just happened to "tickle the sensitive spots" of other truly adaptive functions, such as the rhythmic bodily movements of walking and running, the natural cadences of speech, and the brain's ability to make sense of a cacophony of sounds. Music, Pinker maintained, is what the late paleontologist Stephen Jay Gould called a "spandrel," after the highly decorative but nonfunctional spaces left by arches in Gothic buildings.

Scientific bonding. Researchers at a meeting danced and played in step.

But many researchers disagree, arguing that music clearly had an evolutionary role. They point to music's universality and the ability of very young infants to respond strongly to it as evidence that music itself is hardwired into our brains. "A predisposition to engage in musiclike activities seems to be part of our biological heritage," says Ian Cross, a psychologist of music at Cambridge University. He and others point to the work of University of Montreal neuroscientist Isabelle Peretz, whose studies of musically challenged neural patients, which suggest that distinct regions of the brain specialize in music processing, have made her a leading opponent of the Pinker viewpoint (*Science*, 1 June 2001, p. 1636). Indeed, Cambridge University anthropologist Robert Foley argues that the evidence is suggestive enough that "an adaptive model for music should be the default hypothesis."

All the same, many researchers agree that Pinker's argument represents the key challenge to be met: If music is the result of Darwinian natural selection, how did it evolve, and in what way did it make humans more fit? At the interdisciplinary meeting, many talks focused on music's ability to cement social bonds. Some researchers argued that the roots of music could perhaps be traced back to "performance spaces" created by earlier species of human. Others see music as a way of getting high with one's peers, again to lubricate human bonding. And new studies focus attention on mothers and infants, suggesting that music might have evolved as a way for parents to soothe babies while foraging for food.

By the end of the meeting, says Peretz, "I felt a consensus around the idea that music is not only distinct from language but also has biological foundations." Yet there was also broad agreement that Pinker's challenge had not been fully answered.

* European Science Foundation Workshop on Music, Language, and Human Evolution, Reading, U.K., 28 September to 1 October 2004.

CREDITS (TOP TO BOTTOM): GETTY IMAGES; M. BALTER

Published by AAAS

Sociability versus sex

Like language, most musical behavior leaves no trace in the archaeological record. The earliest undisputed instruments are flutes made from bird bones found at Geissenklösterle cave in Germany and Isturitz cave in France, created and played by modern humans a scant 32,000 years ago. But the first instruments were probably made of perishable materials such as bark or bamboo and are not preserved, says Bannan. And given the universality of music today, most researchers assume that its origins extend back much further, possibly even before modern humans arose some 150,000 years ago. "If there is a strong genetic basis to musicality, then for it to be universally present in the human population it must have been in place more than 150,000 years ago," says Foley.

In the workshop's opening talk, Foley pointed out that Charles Darwin himself was hard put to explain how music made humans better adapted to their environment. In the end, Darwin concluded that music was the result of "sexual selection," the elaboration of traits—such as the peacock's tail—designed to attract a mate and thus ensure reproductive success. Just as some songbirds sing as part of the courtship process, Darwin proposed that humans evolved the ability to sing to each other to express emotions such as love and jealousy.

That theory has some leading proponents today, including University of New Mexico evolutionary psychologist Geoffrey Miller, author of *The Mating Mind*. Miller notes that in some bird species, such as marsh warblers and nightingales, the male signals his supposed genetic fitness to the female by the sheer number of songs he can sing and can reach a repertoire of more than 1000 numbers. He argues that music might have evolved as a way for humans to show off their reproductive fitness. But the sexual selection hypothesis continues to be a minority view among music evolution researchers. "If it was sexual selection, [music] would be a lot more restricted," says Foley. "We would see it more in courtship and less in other activities. Musical ability and activity are too widespread."

Foley and others favor another hypothesis, which holds that in humans, music plays an important role in maintaining social cohesion—critical to mounting coordinated actions—which was essential for hominid survival. Experts in primate behavior have long assumed that cooperation among members of a group boosted the survival rates of early hominids and their offspring, thus selecting for genes that enhance social bonding. But direct evidence has been lacking—until last year, when anthropologist Joan Silk of the University of California, Los Angeles, and her co-workers published a study in *Science*. After 16 years of observing wild baboons, they demonstrated that infants of more sociable female baboons had a higher survival rate (*Science*, 14 November 2003, p. 1231).

Foley points out that the apparent fitness benefit of social cohesion is also the current leading hypothesis for why language itself evolved. "So it makes sense to extend it to music and indeed most other activities," he says. The evening performance led by Espi-Sanchis was a good example of music's "ability to be used in group bonding," adds psychologist Helen Keenoo of the Open University in Milton Keynes, U.K. "Many people seemed to come away from this experience on an emotional high."

First flutes. These 32,000-year-old flutes are the oldest undisputed evidence of music.

CREDIT: FRANCESCO D'ERRICO

But others, including Pinker, say the social-cohesion hypothesis suffers from circular reasoning. Björn Merker, an expert in animal vocalizations at Uppsala University in Sweden who attended the meeting, says that the hypothesis "takes for granted that which it needs to prove, namely why music is needed for bonding and where it got its group-stimulating powers." Merker prefers a hypothesis that "is driven exclusively by the individual advantage of sexual selection." Pinker, who was not at the meeting and is now at Harvard University, adds that "universality and early development don't show that music is an adaptation. It just shows that music is innate. That's a necessary condition for something being an adaptation but not a sufficient one."

Music for the masses

For social-cohesion theorists, the challenge is to explain why singing or dancing enhanced social bonding—and why that in turn fosters greater fitness and survival. Robin Dunbar, a psychologist at the University of Liverpool, U.K., has suggested that music might have put groups of hominids into a collective endorphin high, making them feel more positively disposed toward their fellow hominids—and thus more likely to cooperate and survive. Researchers have long known that listening to music can trigger the production of endorphins, natural opiates that are produced in response to pain or other stress. In a frequently cited 1980 study by Stanford University neuroscientist Avram Goldstein, volunteers who received injections of an endorphin-receptor blocker reported getting considerably less pleasure when they listened to normally moving musical pieces.

Dunbar is well known for his "social brain" hypothesis of human evolution, which holds that larger hominid brain sizes and language both evolved as a response to increasing group sizes in our primate ancestors (*Science*, 14 November 2003, p. 1160). He argues that the endorphin release from music may enhance the subjective feeling of bonding, creating stronger social cohesion. He told the attendees of the meeting about a pilot study that he and his students recently carried out in English churches. In the study, which aimed to look at the effects of music in a social setting, the endorphin levels of churchgoers who attended Anglican services with and without singing were monitored by indirect methods that measured tolerance to pain. (Measuring endorphins directly requires an invasive lumbar puncture.) After services, parishioners who had sung were able to endure having a fully inflated blood pressure cuff on their arms for significantly longer than those who had not sung.

Dunbar stressed that although his own study is very preliminary, the overall evidence suggests that group singing and dancing might have helped bridge what he calls the "endorphin gap" between the nonverbal grooming activities of our primate ancestors and the later development

of language. A number of studies have shown that grooming, which is the social glue of monkeys and many other primates, raises endorphin levels. "Humans are good at finding things that trigger the sensations they like," Dunbar says. And in a social context, he says, "endorphin surges create a very strong sense of bondedness and belonging that seems difficult to create any other way."

One way to support the social-cohesion hypothesis might be to find archaeological evidence of such group interactions in humans' evolutionary past, but such evidence has been hard to come by. In an imaginative talk, archaeologist Clive Gamble of the University of London proposed that group singing and dancing might be traceable back as far as half a million years ago, by seeking evidence for "performance spaces" where such activities might have taken place.

He drew on a recent visit to a village of the Makuri people of northern Namibia, where he watched a performance in which women sat around a fire while men, wearing rattles on their legs and striking sticks, danced around them. The next morning, Gamble could see the circle in the sand made by the male dancers. He compared those circles to several circles, 8 meters in diameter and marked by anvils of bone and stone, unearthed at the 400,000-year-old hominid site of Bilzingsleben in Germany, which he suggested represented gathering and performance areas of these early humans. He also pointed to an unusual concentration of 321 hand axes, many of them unused and all located far from a butchering area, at the 500,000-year-old site of Boxgrove, in Sussex, U.K. Gamble suggested that this possibly symbolic deposit of hand axes may have represented a space where early humans gathered to sing and dance.

Although Gamble's evidence is scant, "I am sure that the hominids at Boxgrove were communicating in a musical and dancelike fashion," says Mithen, who feels that such speculations "give us a perceptive understanding of [early humans'] lifestyle."

Music and motherese

If music did evolve to facilitate a sense of belonging among early hominids, it's possible that a very specific human relationship—that of mothers and infants—was involved, says University of Toronto psychologist Sandra Trehub. She suggested at the meeting that music was crucial to both bonding with and soothing babies, as well as allowing mothers to get on with other tasks that boosted survival.

For years Trehub and her colleagues have studied how mothers talk and sing to their infants. Maternal speech has a number of features that can be considered musical, including higher pitch than normal speech—which is associated emotionally with happiness—and a slower tempo, which is associated with tenderness. Trehub and others have demonstrated that infants prefer maternal cooing to normal adult speech in studies that monitor "infant gaze," or how long a baby spends looking in one direction, considered a measure of attention.

Music to his ears. A mother's song captures her baby's attention.

In a more recent study, in collaboration with Takayuki Nakata of the Nagasaki Junshin Catholic University in Japan, Trehub measured the responses of 6-month-old infants as they watched videos of their mothers. Infant gaze times were even longer during episodes of maternal singing than during normally melodic maternal speech. In another recent study, Trehub and Nakata asked volunteer mothers to talk to their infants for 2 minutes at a time. During one session, the mothers were allowed to touch their babies as much as they wanted; in a second session, they were told not to touch their babies. Trehub and Nakata found that the women markedly increased the pitch of their voices—that is, made them much more musical—when they could not touch their infants. The infants, for their part, responded to their mothers' efforts to compensate for the no-touch rule with even longer gaze times.

Trehub and her co-workers did not try to measure endorphin levels in their infant subjects, but they did measure the cortisol levels in the saliva of babies before and after their mothers spoke or sang to them. Higher blood cortisol levels are a reliable indicator of higher arousal levels, and the hormone passes easily from the bloodstream to saliva. Mothers themselves took the saliva samples by gently swabbing their infants' mouths with a cotton roll. The results were striking: Maternal singing caused a marked decrease in cortisol levels that was maintained for at least 25 minutes after the singing stopped. Maternal speech, on the other hand, caused an initial drop in cortisol levels, which then quickly rebounded to normal. "The function of maternal singing seems to be to regulate the arousal level of the infant," Trehub concluded.

Of course, this is rather obvious to anyone who has ever sung a baby to sleep. But for Trehub, that's the whole point. "Every culture in the world has lullabies," she told the meeting. "And they sound very similar across cultures. They are emotive: The pitch goes up and the tempo goes down." The universality of lullabies, Trehub said, is strong evidence that they have an evolutionary origin. As for what their adaptive function might be, Trehub favors a speculative new idea, called the "putting down the baby hypothesis," recently proposed by anthropologist Dean Falk of Florida State University in Tallahassee.

Falk's hypothesis, in press at the journal *Behavioral and Brain Sciences*, is based on comparisons of the mother-infant interactions of chimpanzees and modern humans as well as data from fossils. She argues that as the brain size of early hominids increased—thus making it more difficult for infant heads to pass through the birth canal—natural selection favored females who gave birth to more immature infants. Unlike baby chimps, who can cling to their mothers at a very young age, human infants are too helpless to do so. The hominid female responded to this situation, Falk argues, by developing melodious vocalizations, or "motherese," so that she could calm and reassure her baby, if not actually put it to sleep, while foraging for food. These vocalizations, Falk concludes, were the prelinguistic forerunner to true language. And although Falk's hypothesis is controversial—not everyone agrees that "motherese" is universal—Trehub says that it is consistent with the notion that maternal singing, and thus early forms of music, also had an adaptive function.

Despite this range of suggestions for music's adaptive functions, Pinker, for one, says his challenge has not been met. "The idea that music evolved to soothe babies might explain why mothers sing to their babies," he says, "but it doesn't explain why older children and adults listen to music." But he adds that whether music was essential to the survival of modern humans has little bearing on its value to us today: "Some of the things that make life most worth living are not biological adaptations."

—MICHAEL BALTER

CREDIT: JOHN CARTER/PHOTO RESEARCHERS INC.

Immunizing Kids Against Flu May Prevent Deaths Among the Elderly

Increasing evidence suggests that vaccinating schoolchildren can create a "herd immunity" that indirectly benefits the unvaccinated

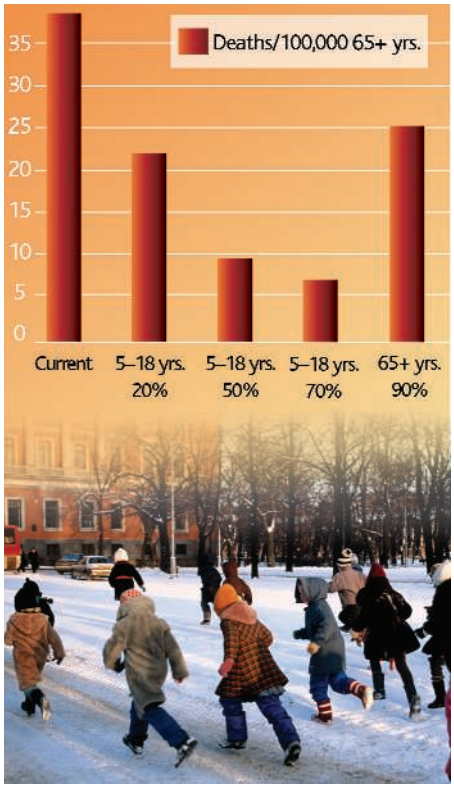
The flu vaccine shortage in the United States has had one clear benefit: It has forced a debate about the best vaccination strategy. And mounting data suggest that there are much more effective ways to combat the annual onslaught of this deadly disease than what the country does now.

The Advisory Committee for Immunization Practices (ACIP) of the Centers for Disease Control and Prevention recommends the vaccine for healthy infants and people 50 or older, as well as people with chronic illnesses, the groups that suffer the most hospitalizations and deaths. But a study published online 2 November in *Vaccine* concludes that vaccinating school-age children could have a greater impact on slowing the spread of influenza virus and reducing disease.

Vaccinating a high percentage of people in a community can decrease the spread of a pathogen and create a "herd protection" that extends to the unvaccinated. Just such a phenomenon has occurred in two neighboring Texas towns, Temple and Belton, report Pedro Piedra, W. Paul Glezen, and colleagues from Baylor College of Medicine in Houston. Since the 1998–99 flu season, the researchers have offered a new, nasally administered flu vaccine to everyone between 18 months and 18 years of age, reaching 20% to 25% of the 20,000 eligible children each year. The researchers tallied serious flu-related disease in all age groups. For comparison, they analyzed three other communities in which less than 1% of the 39,000 children received flu vaccine.

In Temple and Belton, the researchers found that each year they vaccinated school-age children, serious flu cases in adults 35 and older were 8% to 18% lower than in the comparison communities. "This translates to a major reduction in illness," says Piedra. "With the current policy, you only try to control mortality. If you want to control flu, our hypothesis is to focus on kids." Earlier studies showed similar herd protection by vaccinating school-age children, but they did not persuade policymakers. U.S. flu-control campaigns "have never really taken this approach seriously," says epidemiologist Arnold Monto of the University of Michigan, Ann Arbor.

In 1968–69 Monto and colleagues vaccinated 85% of the schoolchildren in the small town of Tecumseh, Michigan. Virtually no adults received the vaccine. Even so, disease rates in all age groups were three times lower in Tecumseh than in a largely unvaccinated nearby town. "There's a lot of benefit to be gained by targeting school-age kids," says Monto, who chairs a subcommittee of ACIP that focuses on herd immunity. Monto adds that vaccinating the "frail" elderly may be less effective because they often do not develop a robust immune response.



Kid you not. Vaccinating even 20% of school-age children may prevent elderly deaths from flu more effectively than increasing elderly vaccination rates.

Still more evidence of a herd effect comes from Japan. As Thomas Reichert, Baylor's Glezen, and co-authors reported in the 22 March 2001 *New England Journal of Medicine*, Japan vaccinated 50% to 85% of schoolchildren during the mid-1970s and 1980s, but the elderly rarely received the vaccine. During that period, deaths from influenza and pneumonia—which mainly kill the elderly—dropped by at least 10,000 per year. As of 1987, parents could exempt their children from the program; deaths from those diseases began to steadily increase.

Biostatisticians Ira Longini and Elizabeth Halloran of Emory University in Atlanta, Georgia, have developed a mathematical model to find the optimal way to distribute flu vaccine. As they describe in a 2000 *Vaccine* paper, vaccinating just 30% of schoolchildren in a community reduces the likelihood of epidemic spread of flu from 90% to 65%. If half the children are vaccinated, the likelihood drops to 36%; if 70% are vaccinated, the probability of epidemic spread plummets to 4%. "Children are highly connected among themselves, and they're connected through adults to families and neighborhoods," says Longini, noting that they're twice as likely to become infected as adults. "By vaccinating children, you're tearing the heart out of that web."

As was found in the Japanese study, Longini and Halloran's models suggest that such a strategy could greatly reduce deaths among the elderly. If, for example, coverage of schoolchildren increased from the current 5% to 20%, they predict it would reduce more deaths in the over-65 population than increasing their vaccination coverage from the current 68% to 90% (see graph).

Given the evidence, why hasn't herd immunity to flu received more attention? "Flu until now wasn't really a sexy topic," says Halloran. "And it's sort of a medical thing that you look at protecting people directly."

Although the results hold for both the injected vaccine and the recently licensed inhaled one, Monto suspects that the latter works better. The nasal flu vaccine relies on live but weakened virus, whereas the injected version contains killed influenza. In theory, the live vaccine can trigger a broader immune response and may better thwart viral transmission because it stimulates immunity at the site where the virus typically enters. For mass immunization campaigns, says Monto, the spray is also easier to deliver.

The Texas herd-immunity study will continue for two more seasons, and MedImmune Inc. in Gaithersburg, Maryland, the maker of the nasal vaccine, has four of its own studies under way that will attempt to assess the indirect benefits of vaccinating schoolchildren.

In another novel strategy, two reports published online 3 November by the *New England Journal of Medicine* suggest that the current vaccine supply can be "stretched" by injecting smaller doses under the skin, instead of intramuscularly. This approach offers "great promise," wrote Anthony Fauci and the late John La Montagne of the National Institute of Allergy and Infectious Diseases in an accompanying editorial. "It's not going to be a practical solution to the shortage problem we have this year," stresses Fauci, "but it could be in the future." **–JON COHEN**

CREDITS: (GRAPH) IRA LONGINI; (PHOTO) DEMETRIO CARRASCO/GETTY IMAGES

Published by AAAS

RNAi Shows Cracks in Its Armor

RNAi's tendency to influence genes and proteins that it's not designed to target is provoking questions and controversy, as scientists labor to solve the problem

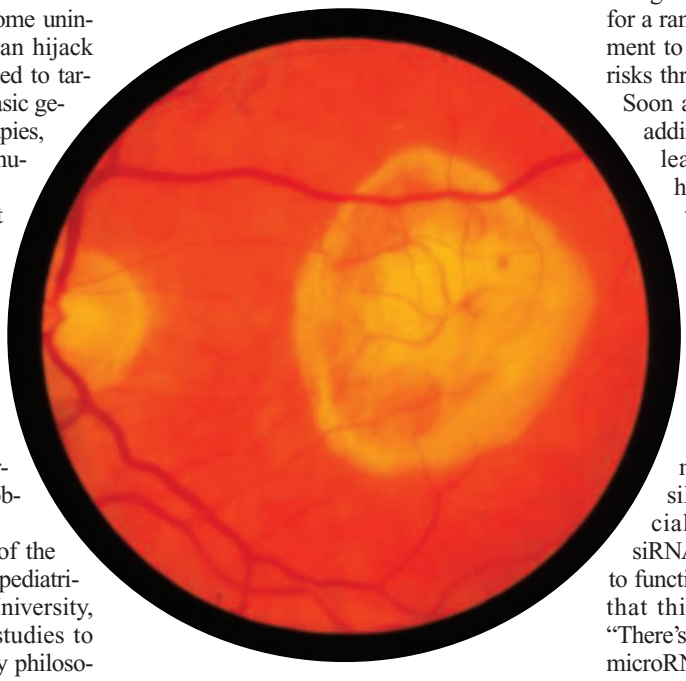
A promising new approach to manipulating genes is showing blemishes as it moves from its glamorous early days to a more nuanced adolescence. The technique, RNA interference (RNAi), shuts down genes; this braking effect helps reveal a gene's function and could potentially be used to treat a host of diseases. But a growing number of researchers are learning that RNAi, which was hailed for its laserlike specificity by scientists and the press (including *Science*, which anointed it 2002's Breakthrough of the Year), comes with some unintended baggage. In particular, it can hijack genes and proteins it wasn't designed to target—a potential problem for both basic genetics studies and RNAi-based therapies, some of which are just beginning human testing.

Even experts concerned about these so-called off-target effects hasten to point out that RNAi's future remains bright. But the issue is stirring controversy in the field. Biologists are struggling to determine—and agree upon—just how widespread off-target effects are, why they occur, and what can be done to avoid them. Some are feverishly working to circumvent the problem, with early hints of success.

"We don't know all the rules" of the RNAi machinery, says Mark Kay, a pediatrician and geneticist at Stanford University, who's conducting RNAi animal studies to treat hepatitis B and C viruses. "My philosophy is that we move forward with caution, but we move forward."

In the late 1990s, scientists discovered the potency of small RNA molecules just 21 nucleotides or so in length—some lab-produced, others naturally occurring. Injecting these RNAs, often called small interfering RNAs (siRNAs), into worms and flies silenced only messenger RNA (mRNA) molecules containing a complementary sequence. That, in turn, blunted expression of the gene producing that messenger RNA. In these organisms, there was no sign that an mRNA with a slightly mismatched sequence—with, say, 17 compatible nucleotides out of 21 in the siRNA—could also be affected.

But as scientists moved on to studies in mammals, the picture changed. One of the first to see irregularities was Peter Linsley, the executive director of cancer biology for Rosetta Inpharmatics in Seattle, Washington, a subsidiary of the drug giant Merck. "We thought it would be cool," Linsley recalls, to use siRNAs to try to design more targeted drugs. The plan: Use siRNAs to knock down expression of a particular gene that an experimental compound is already designed to target. Then add that compound to the mix, and see if it disrupts other genes as well—something that might suggest it's not targeted enough for treating patients.



Eyeing RNAi's potential. With RNAi trials launching for macular degeneration (above), researchers are watching closely to see whether the technique has any unexpected effects on humans.

But as it turned out, says Linsley, it wasn't the compounds that were poorly targeted. The siRNAs were turning down expression in multiple genes instead of just one. "The siRNAs were dirtier than our compounds," says Linsley, whose team was taken aback. The pattern persisted, and the researchers finally concluded that siRNAs could "cross-react" with other genetic targets. After some struggle convincing reviewers that the paper was accurate, it appeared in *Nature Biotechnology* in June 2003.

RNAi enthusiasts responded skeptically. After all, they'd trusted for several years that the small RNA molecules they were crafting were undeniably specific. Gradually, prodded by Linsley's work and in some cases their own, that belief shifted. "We saw more and more unexplained phenomena," says René Bernards, a cancer geneticist at the Netherlands Cancer Institute in Amsterdam. Phillip Zamore, a biochemist at the University of Massachusetts Medical School in Worcester, says his thinking evolved "when I couldn't find a way to disprove Peter Linsley." Like many of his colleagues, Zamore now believes that RNAi's limitations should have been obvious and that to presume such specificity was "incredibly unreasonable." Genetics, says Zamore, is rarely so neat.

Why off-target effects occur remains a matter of debate. One possibility is that introducing foreign siRNAs into a cell's existing RNAi system—upon which it relies for a range of functions, from early development to protecting the genome's integrity—risks throwing a wrench into the machinery. Soon after scientists began experimentally adding siRNAs to mammalian cells, they learned that these cells naturally use hundreds of so-called microRNAs, which are similarly sized and help translate RNA molecules into proteins. MicroRNAs are widely considered much less specific than siRNAs, frequently targeting sequences that only partly match their own.

This has left scientists wondering whether mammalian cells, awash in microRNAs, are mistaking foreign siRNAs for more of the same, especially because both microRNAs and siRNAs need many of the same enzymes to function. An RNAi study last year showed that this mistaken identity could occur. "There's probably a fine balance between the microRNA pathway and what we're putting into cells of animals," says John Rossi, a molecular biologist at the City of Hope Graduate School of Biological Sciences in Duarte, California.

Weak sequence matching between siRNAs and genes has also been traced to a specific part of the siRNA. That bit, called the 5′ end, helps govern how an siRNA binds to its target. As Linsley found and others such as Zamore confirmed, if that particular piece, about seven nucleotides long, matches a sequence in another gene, there's a risk of the entire siRNA binding to that gene instead.

Increasingly, biologists are turning up other seemingly esoteric details that may also determine whether an siRNA shuts down unintended genes. In the fall of 2003, a group led by Anastasia Khvorova at Dharmacon, a company in Lafayette, Colorado, and another led by Zamore, reported in *Cell* that siRNAs with certain sequences and structures unwind

CREDIT: PAUL PARKER/PHOTO RESEARCHERS INC.

Published by AAAS

slightly differently—and the pattern in which they unwind can ultimately affect how good they are at targeting the right gene.

A year ago Bryan Williams, a cancer biologist at the Cleveland Clinic in Ohio, identified another, more controversial kind of off-target effect. In fruit fly cells and human cancer cells, he found that siRNAs activated the interferon pathway, which is the body's first line of defense against viruses.

How widespread the interferon response is remains uncertain. Although some unpublished reports of interferon response in animals exist, "by and large, people who've treated animals with siRNAs have not seen significant interferon induction," says Phillip Sharp, a biologist at the Massachusetts Institute of Technology in Cambridge and co-founder of the company Alnylam, which next year hopes to begin testing RNAi in patients with the eye disease macular degeneration.

On target? Hepatitis B is one of the diseases RNAi enthusiasts are working to disable.

In animals generally, the impact of off-target effects isn't clear. Mice with liver disease treated with RNAi technology do suffer toxic effects, says Harvard's Judy Lieberman, but those are considered more a result of the way siRNAs are delivered—in this case, under extremely high pressure, to ensure that they infiltrate liver cells. Lieberman says she's seen no visible evidence of off-target effects in her mice, but she is planning to examine the animals more carefully. Says Linsley, "You can't conclude it's not there until you look."

Looking, though, can be trickier than it sounds. For the most part, scientists are relying on microarrays, which show gene-expression levels, to learn whether their siRNAs are hitting unintended genes; in general they're finding that a dozen genes may be affected by a single siRNA. (Linsley has recorded on average at least 40.) Still, it's difficult to gauge how big a problem that is. Mismatches provoke a less dramatic change in gene expression than complete matches. Most gene expression varies by less than twofold when the siRNA doesn't fully match—often not enough to have a substantial biological impact on how a cell, or an animal, actually functions.

Cautious but upbeat. Stanford's Mark Kay hopes that off-target effects won't derail RNAi's extraordinary possibility.

CREDITS (TOP TO BOTTOM): CDC/PETER ARNOLD INC.; MARK KAY/STANFORD UNIVERSITY

But using microarrays to look for off-target effects has one big drawback: They show only gene expression, not protein levels. If siRNAs are imitating microRNAs, that means they're not affecting DNA directly but rather are altering how RNA is translated into protein. Microarrays thus might not detect changes in protein abundance. "What's really important is what's happening at the protein level, and we don't have a lot of data on that," says Linsley.

Researchers at Dharmacon and elsewhere are trying to see whether microarray results correlate with changes in protein levels. At a meeting last week in Titisee, Germany, Sharp presented preliminary data from his lab showing a 10-fold change in protein levels with only a twofold microRNA difference, the level commonly seen from an off-target effect. But doing the kind of broad protein screens that microarrays today accomplish for genes isn't yet possible. "You can spend the rest of your life trying to see all 10,000 proteins in the cell," says Sharp, "and you'll never get an answer."

Scientists are quick to add a caveat: Even if off-target effects occur, they don't necessarily affect the phenotype, or how a cell or animal actually functions. If phenotype isn't altered, notes Zamore, the effects rarely make a difference.

The significance of off-target effects also depends on how RNAi is used: to unearth the function of mystery genes or as a medical therapy. In the first case, scientists are getting around the problem by applying several different siRNAs, each of which corresponds to a different sequence in their gene of interest. That way, if one siRNA prompts an off-target effect that changes a cell's phenotype, it will be more apparent.

When it comes to RNAi-based treatment, though, the potential challenges multiply. The first clinical trial of RNAi therapy—for use in macular degeneration—was launched last month by the Philadelphia company Acuity Pharmaceuticals. Because treatments can be restricted to the eye, the risk of off-target effects is of less concern.

For other diseases, "it's unclear how much of an issue this is going to be," says Stanford's Kay, whose RNAi work focuses on hepatitis. The disease is a popular choice for RNAi therapies because RNAi can disable the virus. Yet it's also difficult to target the liver without affecting other parts of the body. In Kay's view, RNAi therapies shouldn't be viewed differently from traditional drugs: "If you give somebody aspirin, they're going to have changes in gene expression in specific tissues." He expects that RNAi clinical trials, like all others, will need to home in on the lowest effective dose and monitor patient safety carefully.

To avoid any effects that may cause problems, researchers are chemically modifying siRNAs to try to stop them from glomming onto messenger RNAs they should ignore. Modifications can also make the key 5′ bit of siRNAs more sluggish in its binding, rendering mismatches less likely. Linsley and some Dharmacon colleagues have just submitted a paper on the subject. "We've made a few steps," he says, declining to be more specific. But "I don't think we've completely solved it." Although off-target effects may forever linger as a risk of RNAi, he and others say, they hope that they'll become less of a worry, and soon.

–JENNIFER COUZIN

SAN DIEGO—From 23 to 27 October, this California coastal city hosted the annual Society for Neuroscience meeting, at which more than 30,000 researchers presented data on topics such as sleep problems, addictive anesthesia, and baby talk.

Brain Cells May Pay the Price for A Bad Night's Sleep

It's enough to keep you up at night: Sleep apnea, a condition in which breathing irregularities occur during sleep, may kill neurons in brain regions crucial for learning and memory, according to research on rodents described at the meeting. The findings may provide a disturbing explanation for the cognitive deficits often seen in people with sleep apnea. And to make matters worse, new evidence suggests that adding an unhealthy diet to the mix greatly compounds the neural harm caused by disordered sleep.

You snooze, you lose. Rodent research suggests that sleep apnea can kill brain cells.

In the United States, sleep apnea affects at least 2% of children, 4% of middle-age adults, and 10% of older adults, according to conservative estimates, and it is even more common in obese people. The cognitive problems that result, including hyperactivity, wandering attention, and learning deficits, were long thought to stem solely from the fatigue that follows a bad night's sleep, says Gordon Mitchell, a respiratory neurobiologist at the University of Wisconsin, Madison. "The concept that you're causing specific damage through cell death in particular areas of the brain is pretty new," he says.

Mitchell and others credit their colleague David Gozal with much of the work that has led to the unanticipated realization. Gozal, a pediatric researcher at the University of Louisville in Kentucky, has published a slew of papers in the last few years that describe what happens in the brains of rodents exposed to brief periods of reduced oxygen similar in nature to those experienced by a person with sleep apnea.

It's not a pretty picture. Gozal's team has found that intermittent hypoxia kills rodent brain cells in the hippocampus, a key memory center, and interferes with a process called long-term potentiation, a strengthening of neural connections considered crucial for learning and memory. The same reduced level of oxygen, kept constant, has little or no effect.

Gozal has also begun to elucidate some of the molecular events underlying the damage. Intermittent hypoxia stresses brain cells, causing them to produce molecules called oxygen free radicals that are notorious for wreaking havoc on cells and driving them to self-destruct. Inhibiting certain enzymes involved in the stress response can save neurons and prevent learning deficits in rodents subjected to hypoxic periods, Gozal's team has found. They also reported at the meeting, for the first time, that regular exercise—the rat equivalent of walking in the park for an hour a day, Gozal says—cancels the learning deficits caused by intermittent hypoxia.

That's the good news. The bad news is that a diet high in fat and refined carbohydrates appears to magnify the deleterious effects of intermittent hypoxia. The diet alone caused a mild learning impairment when rats were tested in a water maze, and it reduced the level of the activated form of a protein called CREB in their hippocampi. CREB plays an important role in memory consolidation and neuron survival, so it is a good general marker of hippocampal health, Gozal says. Pairing the high-fat, refined-carbohydrate diet with intermittent hypoxia had a synergistic negative effect on learning and activated CREB. "It's a major disaster for the brain," Gozal says. If the finding holds true for humans, he adds, that would be especially troubling because sleep apnea and unhealthy diets are a common combo for many people with obesity.

"It's an incredibly important observation," says Sigrid Veasey, a neuroscientist at the University of Pennsylvania in Philadelphia. "The injuries can be compounded by what would be considered an unhealthy diet, but [is] probably a pretty standard diet for a lot of people in this country."

—GREG MILLER

Anesthesia's Addiction Problem

Warning: Anesthesiology may be hazardous to your health. According to Mark Gold, who has spent much of his career investigating the problem of physicians who abuse drugs, data from the state of Florida show that "every year since 1995, anesthesiologists were the number one [medical] specialty for substance abuse or dependence." His group found, for example, that in 2003, anesthesiologists represented less than 6% of all physicians in the state but made up almost 25% of the physicians monitored for substance-abuse disorders.

Although Gold, chief of the McKnight Brain Institute at the University of Florida (UF), isn't the first to conclude that anesthesiologists are especially susceptible to abusing drugs, particularly opiate-based compounds similar to those used in general anesthesia, he has a new explanation. The problem isn't simply easy access to drugs, he says. Instead, he and his colleagues propose that the physicians may become primed for drug abuse because they chronically inhale small amounts of anesthetics that sensitize the brain's reward pathways. Indeed, at the meeting, Gold's team reported finding traces of intravenously delivered anesthetics in the air of operating rooms.

When early anesthesiologists depended on gases such as ether and chloroform, secondhand exposure was a serious problem. But better ventilation and increased use of intravenous drugs reduced such exposures. As for the issue of drug access, hospitals have gone to great pains to safeguard their medications. Even so, anesthesiologists continue to have much higher rates of opiate-related substance abuse than other physicians with similar access.

From his review of records from Flori-

CREDIT: CORBIS

Published by AAAS

Breathing problem. IV-delivered anesthetics may escape into the air and affect physicians.

da's Impaired Professional programs, Gold has found that anesthesiologists who abuse drugs tend to start much later in life than other addicts, who typically experiment with drugs during their youth. Anesthesiologists also tend to relapse unless they change professions, says Gold.

Having conducted research on whether secondhand smoke sensitizes brain reward pathways—children of smokers are much more likely to smoke—Gold wondered whether secondhand anesthesia might be at work. He teamed up with several UF anesthesiologists to answer that question. As a first step, they used a mass spectrometer to examine the air in operating rooms during cardiac bypass surgeries in which fentanyl, an opiate many times more potent than morphine, and a nonopiate anesthetic, propofol, were administered intravenously to patients. Both compounds were found in the operating room air and at higher concentrations in the space between the anesthesiologist and patient, Matt Warren, a graduate student with the Florida team, reported in San Diego.

In another test, the researchers found that volunteers given fentanyl exhale it. In lengthy operations such as cardiac bypasses, anesthesiologists "could be breathing analgesics and anesthetics for 8 hours," speculates Gold.

The Florida team is now collecting blood samples of anesthesiologists during operations to see if fentanyl is present. They also intend to expose rodents to the same air concentrations of the anesthetics as found in the operating rooms and test whether those animals are more susceptible to drug addiction and develop changes in brain regions involved in the rewarding aspects of drugs.

Those follow-up studies will be needed to convince anesthesiologists and others who are intrigued by Gold's hypothesis. "He's a guy who thinks outside the box," says Robert L. Dupont, president of the Institute for Behavior and Health in Rockville, Maryland, and former director of the National Institute on Drug Abuse. "But it's hard for me to imagine that the doses people are getting this way are having any biological effect. I'm ready to be persuaded, but I'm skeptical."

–JOHN TRAVIS

Listen, Baby

How quickly babies home in on the sounds of their native language during their first year may predict how quickly they learn new words, string together complex sentences, and acquire other language skills as toddlers. The new research, presented in San Diego, helps pin down a milestone in language development and may shed light on why the ability to pick up a new language wanes with age.

When it comes to language, babies are "citizens of the world," Patricia Kuhl of the University of Washington, Seattle, said in a lecture here. In the early 1990s, her team found that 6-month-old infants naturally possess a language skill far beyond the reach of adults: They can distinguish all the sounds of all the world's languages—about 600 consonants and 200 vowels. By the end of their first year, however, babies begin to specialize. As they become better at recognizing the basic elements, or phonemes, of their native language in all their acoustic variations—learning to lump /o/ as pronounced by mom together with /o/ as pronounced by grandpa and Bugs Bunny, they lose the ability to distinguish phonemes in other languages. By about 11 months, for example, certain vowels that sound distinct to Swedes start to sound the same for a baby born to an English-speaking family.

Baby talk. An infant's level of speech discrimination predicts language skills.

Kuhl and others have argued that this change in speech perception is an essential step in language learning. They contend that babies need to be adept at identifying native phonemes, for example, before they can break down a stream of speech into individual words. That skill, in turn, is necessary for assigning words to objects, creating more complex sentences, and so on.

Recent work in Kuhl's lab bears this out. Her team has been using electroencephalogram (EEG) electrodes to monitor the brain activity of 7-month-old infants, who are just at the cusp of the change in phoneme perception. The babies listened to a recorded voice repeat a single phoneme several times before switching to another phoneme. If the baby caught the switch, a blip appeared in the EEG record. This evoked related potential (ERP) is a standard indicator that the brain has picked up something new. The researchers tested the babies' ability to discriminate both native and non-native phonemes and then followed up with a battery of language tests at 14, 18, 24, and 30 months of age.

The ERP recordings revealed that infants who at 7 months of age were good at native phoneme discrimination tended to be bad at non-native phoneme discrimination, and vice versa. This fits with the "neural commitment" theory proposed by Kuhl several years ago, says Mirella Dapretto of the University of California, Los Angeles. Kuhl's work suggests that "the more your brain gets committed to picking up what's relevant in the first language you're exposed to, the more you're tuning out distinctions that are relevant in other languages," Dapretto says.

The follow-up studies suggest that the brain's commitment to native speech sounds provides the foundation for later language learning. Although all the children in Kuhl's study tested in the normal range, the ones who did best at native phoneme discrimination at 7 months scored higher at later times on all language measures, including number of words produced, duration of utterances, and sentence complexity.

"It's extremely interesting [that] you can look at infants and learn something really important about their future learning," says April Benasich, a cognitive neuroscientist at Rutgers University in Newark, New Jersey. Benasich says the findings add to evidence, including work from her own lab, that it may be possible to screen young infants to identify those likely to need extra help with language learning.

–GREG MILLER

CREDITS: GETTY IMAGES

Published by AAAS

RANDOM SAMPLES

Edited by Constance Holden

Tiwanaku goblet

Life in the Ancient Andes

Archaeologists have discovered a 1000-year-old trove of finely painted and beautifully sculpted drinking vessels and figurines on an island in Bolivia's Lake Titicaca. Some of the jugs, 20 of which are unbroken, are shaped like human heads and animals, offering a rare glimpse of the Tiwanaku culture.

The finds, announced last month by Antti Korpisaari and Martti Pärssinen of the University of Helsinki working with Bolivian colleagues, also included the shards of hundreds of vessels that priests may have ceremonially tossed into a pit. The workmanship represents "some of the very highest artistic achievements of Tiwanaku potters ever discovered," says Korpisaari.

Because textiles haven't survived in the moist Andean highlands, archaeologists have little information on how the people were clothed. "To actually see what people might have looked like while alive is extremely valuable," says Deborah Eileen Blom of the University of Vermont, Burlington. The ceramics also show jewelry and tattoos.

The Definitive Mozart

Image not available for online use.

A filmmaker hopes to use DNA to settle a long-standing mystery: Does an Austrian foundation really have Wolfgang Amadeus Mozart's skull? Researchers have already discreetly unearthed bones of several Mozart family members for a documentary celebrating the composer's 250th birthday in January 2006. In the coming months, they will attempt to compare DNA samples taken from the bones to a sample from the skull, held by the International Mozarteum Foundation in Salzburg.

The skull was supposedly unearthed by a gravedigger several years after the 35-year-old composer was buried in an unmarked pauper's grave in December 1791. It passed through several hands before arriving at the Mozarteum Foundation in 1902. Decades of historical, dental, and even facial reconstruction studies have been frustratingly inconclusive, says foundation director Stephan Pauly.

Journalist Burgl Czeitschner, who proposed the project, says the team has unearthed bones thought to be those of Mozart's maternal grandmother and his niece—both of whom would carry the same mitochondrial DNA as the composer. If the samples match, Czeitschner says, then scientists might be able to use the skull to learn more about how Mozart lived and the unresolved mystery of how he died. (One persistent rumor has him being poisoned by a jealous rival composer.) Czeitschner says producers plan to keep a tight wrap on any results until the film is broadcast.

Betting on Bush

In a year when pre-election polls fluctuated, election markets (*Science*, 30 July, p. 603) managed to send a relatively consistent message.

From the end of the Republican convention on 2 September through Election Day, the odds on President Bush's eventual victory never dropped below 50% on the commercial Tradesports.com Web site. On election eve, a $1 contract for Bush to win closed at 53 cents on the dollar, meaning he was believed to have a 53% chance of winning.

On the Iowa Electronic Market, run by the University of Iowa, Bush consistently led in the betting. A late Kerry surge brought the Democrat up to only 49.5% in the "vote share" market—a 1.5% overestimate (as calculated for a two-way race), which was nevertheless within the market's historical error rate. And there was no need to sweat all night over swing states Florida and Ohio: Election followers had only to log on to Tradesports.com to see that they were pegged in the GOP camp.

According to Iowa accounting professor Joyce Berg, the Iowa market stayed steadier and closer to the final results throughout the year than polls, which underestimated Bush's final vote share before the Republican convention, then overestimated it until the first debate. Even after the last debate, "the polls were all over the map" compared with the markets, says Berg.

AIDS Art in Paris

This anti-AIDS mural photographed in South Africa is part of a traveling photography exhibition, "Sciences au Sud," focusing on French research efforts in the Southern Hemisphere, which opened in Paris last month.

CREDITS (TOP TO BOTTOM): ANTTI KORPISAARI; ARCHIVO ICONOGRAFICO, S.A./CORBIS; ELIZABETH DELIRY ANTHEAUME/IRD

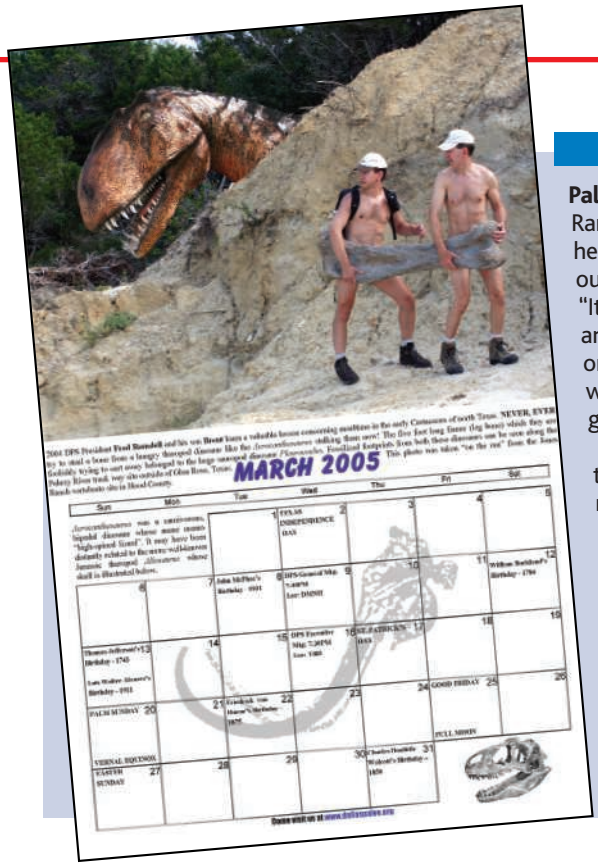
Published by AAAS

Edited by Yudhijit Bhattacharjee

IN BRIEF

GAVI head. British health policy expert Julian Lob-Levyt is the new executive secretary of the Global Alliance for Vaccines and Immunization (GAVI), based in Geneva, Switzerland. He succeeds Tore Godal, who has held the post since the alliance was formed in 2000. GAVI, which is a partnership of government agencies and private foundations worldwide, and its financing arm—The Vaccine Fund—have so far given out $429 million for vaccinations in more than 70 countries. Lob-Levyt has worked at the United Kingdom's Department for International Development since 1998 and is a senior policy adviser to the executive director of the Joint United Nations Programme on HIV/AIDS.

New NPA chair. Keith Micoli, a postdoc in pathology at the University of Alabama, Birmingham, has been named chair of the National Postdoctoral Association for 2005, succeeding founding chair Carol Manahan. The 2-year-old association has also elected five executive board members for 2-year terms starting 2005.



IN PRINT

Paleo buffs. Nothing gets between Fred Ransdell and his fossils—at least not when he's posing as Mr. March for a calendar put out by the Dallas Paleontological Society. "It's supposed to be educational, attractive, and funny," says Ransdell, an amateur paleontologist and president of the society, which hopes the proceeds will finance two graduate scholarships awarded each year.

The calendar features "revealing but tasteful" photos of local fossil enthusiasts, men and women, posing in the field with nothing but a smile and their favorite specimens. For their part, Ransdell and his son Brent are seen toting a cast of the femur of a sauropod called *Pleurocoelus* found at a quarry about 150 kilometers south of Dallas. In the background lurks a predator from that era, 110 million years ago, which is actually the image of a dino sculpture superimposed on the photo.

To order a copy, visit dallaspaleo.org.

DEATHS

Microbe expert. John La Montagne, deputy director of the National Institute of Allergy and Infectious Diseases (NIAID), died on 2 November en route to a Pan American Health Organization meeting in Mexico City. The 61-year-old microbiologist was waiting in a passport line at the Mexico City airport when he collapsed suddenly. NIAID officials say he had no serious health problems, and the cause of death is as yet unknown.

Born in Mexico, La Montagne

worked at NIAID on a wide range of infectious diseases including influenza, malaria, and tuberculosis. "He had an encyclopedic knowledge of microbiology," says Anthony Fauci, La Montagne's boss and longtime friend. "And he had to be one of the finest human beings anyone of us had ever met."

La Montagne is survived by his wife of 35 years.

JOBS

Tackling big questions. University of Chicago cosmologist Edward Kolb has been chosen to head the new Particle Astrophysics Center at Fermi National Accelerator Laboratory in Batavia, Illinois.

The center will bring together some 40 scientists from Fermi's experimental and theoretical astrophysics groups, along with new hires, to focus on the growing overlap between astrophysical observations and high-energy particle physics experiments. It will analyze findings from projects such as the Sloan Digital Sky Survey and the Pierre Auger Cosmic Ray Observatory to investigate questions on dark matter and the big bang.

"There are so many new ideas and so much new data," says Kolb, 53, who came to Fermilab in 1981 and helped found its theoretical astrophysics group. Kolb hopes to attract private funding in addition to in-house support for the new center, whose debut next month features a lecture by Nobelist Riccardo Giacconi and a workshop on galaxy clusters.

Got any tips for this page? E-mail people@aaas.org

CREDITS (TOP TO BOTTOM): NIH/NIAID; FERMILAB/VMS

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 6 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

The Future of NASA

AMONG THIS YEAR'S MOST SIGNIFICANT science findings was the Mars Exploration Rover Opportunity's discovery that Meridiani Planum was once subsumed under an ancient salty sea. Through missions like Cassini-Huygens and Genesis, and the Spitzer Infrared telescope, NASA continues to conduct cutting-edge science. I don't feel *Science*'s readers were told any part of that story in Andrew Lawler's article "Rising cost of shuttle and Hubble could break NASA budget" (News of the Week, 24 Sept., p. 1882).

NASA is studying the many options for a potential robotic servicing mission of the Hubble Space Telescope. The cost estimates mentioned in Lawler's article are premature. We continue to search for the best possible solution in servicing Hubble. It is too early to place a price tag on this mission.

CREDIT: NASA

Hubble Space Telescope

There were other inaccuracies in the article. First, I have not assigned managers in the Science Directorate to "find $400 million in cuts so that the space shuttle could resume flying." On the contrary, NASA is forging ahead, hopeful that Congress will enact the President's full budget request.

Second, I never declined an interview for the article. I said at the time of the interview request that it was a sensitive juncture in the appropriations process and to talk about our budget at that time would have been counterproductive.

Fortunately, NASA has a brighter future than the one portrayed in Lawler's article. By pursuing the goals of our Vision for Space Exploration, we expect to significantly advance our scientific frontiers.

AL DIAZ
Assistant Administrator for Science, National Aeronautics and Space Administration, 300 E Street, SW, Washington, DC 20546–0001, USA.

Research Corporation and John Schaefer

I WAS DISAPPOINTED IN JEFFREY MERVIS'S article "Shooting for the stars" (News Focus, 27 Aug., p. 1231), about Research Corporation (RC) and its retiring president, John P. Schaefer. There are many debatable points in Mervis's piece, but one allegation misleads the scientific community even more than it might damage Research Corporation.

Mervis quotes Joan Valentine as saying "I don't think that John [Schaefer] ever cared very much about the grants programs; his passion was astronomy." That statement is as parochial as it is false. In the 12 years that I have worked closely with Schaefer, he has repeatedly demonstrated his passion for the advancement of all science.

Schaefer has always characterized the grant programs as core to the foundation's mission. From 1992 to 2003, RC spending on these programs had more than doubled and accounted for 68% of total RC expenditures. The community we serve needs to understand that RC is committed to academic research and is ready—as always—to provide funding to faculty doing cutting-edge research in the physical sciences through its regular programs.

RAY KELLMAN
Vice President, Research Corporation, 4703 East Camp Lowell Drive, Suite 201, Tucson, AZ 85712, USA. E-mail: ray@rescorp.org

WE READ WITH INTEREST THE RECENT PROFILE of John P. Schaefer ("Shooting for the stars," J. Mervis, News Focus, 27 Aug., p. 1231), president of the Research Corporation (RC), the oldest scientific research foundation in the country. An objective reading of the article leaves the distinct impression that the grant programs of RC (www.rescorp.org) were taking a back seat to large projects such as the Large Binocular Telescope (LBT). Those quoted in the article suggest that "RC was abandoning its historical mission … of advancing the natural sciences through early-career awards and small grants that foster innovative research and teaching." As current and former members of the RC grants advisory committee, we strongly disagree with these opinions.

The grant programs at RC are stronger than ever and making a vital difference to 150 to 200 annual recipients. Total annual expenditures on grants at RC have steadily increased each year, from $2.5 million in 1981 (the year before Schaefer took over) to over $6.2 million in 2003. RC grant programs include a long-standing program of research support at undergraduate institutions (Cottrell College Science Awards), a new program for beginning faculty who excel at both research and teaching (Cottrell Scholar Awards), and a program for established scientists seeking to explore new research areas or help reestablish federally funded research programs (Research Opportunity Awards).

We have watched Schaefer at close range as he strongly supported and nurtured the RC grant programs. The country's scientific research community has gained consistent, essential grant support from RC over the past 22 years.

DON DEARDORFF,[1] PETER DORHOUT,[2] NANCY HAEGEL,[3] BRENT L. IVERSON,[4] MICHAEL MORRISON,[5] JACK PLADZIEWICZ,[6] MATS SELEN,[7] THOMAS D. TULLIUS,[8] TIMOTHY S. ZWIER[9]
[1]Occidental College, Los Angeles, CA 90041, USA. [2]Colorado State University, Fort Collins, CO 80523, USA. [3]Pacific Grove, CA, USA. [4]University of Texas at Austin, Austin, TX 78712, USA. [5]University of Oklahoma, Norman, OK 73019, USA. [6]University of Wisconsin, Eau Claire, Eau Claire, WI 54702, USA. [7]University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [8]Boston University, Boston, MA 02215, USA. [9]Purdue University, West Lafayette, IN 47907, USA.

The File-Drawer Problem, Revisited

IT HAS BEEN CONTENDED THAT DRUG companies have hidden, in the "file drawer," the results of unsuccessful clinical trials while publishing the results of more successful trials (*1*). As others have pointed out, no one's interest (drug company, doctor, patient) is served by the marketing of ineffective drugs, but nevertheless, the

Published by AAAS

scientific enterprise is set back when full results are not published. The remedy proposed is to make all clinical trials data available ("The old file-drawer problem," D. Kennedy, Editorial, 23 July, p. 451); indeed, some companies are already planning to post their clinical trials data.

If there is minor mischief in the drug company file-drawer, there is major mischief with many epidemiology studies. Although there are legitimate reservations (*2*), why shouldn't epidemiologists also open their data sets (*3*)? Arguably, the scientific need is greater than that for clinical trials. Whereas clinical trials test prespecified, well-defined hypotheses, in epidemiological investigations, hundreds to thousands of questions may be asked, but only those results that pass the minimal threshold of "statistically significant, $P < 0.05$" are reported. Statistical analyses are often not adjusted to reflect multiple testing, even though procedures are readily available (*4*). If hundreds to thousands of results are locked in the epidemiology file drawer, even readers who wish to make the adjustment for themselves cannot. The worst-case scenario is that the vast majority of results reported from epidemiology studies are false positives.

To cite (but not single out) only one example of epidemiology file-drawer cases, antibiotic use was associated with breast cancer (*5*). The antibiotic–breast cancer association was culled from a very large search process. If analyses were adjusted to reflect the number of questions, there might well be no significant association. The central point, however, is that keeping these results in the file drawer prevents others from doing alternative analyses, and again the loser is the scientific enterprise.

S. STANLEY YOUNG[1] AND HEEJUN BANG[2]

[1]National Institute of Statistical Sciences, Research Triangle Park, NC 27709, USA. [2]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA.

References
1. L. Abboud, "Lilly plans broad access to results on its drug trials," *Wall Street J.*, 3 Aug. 2004, pp. B1–B2.
2. E. Marshall, *Science* **290**, 2829 (2000).
3. Board on Life Sciences, *Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences* (National Academies Press, Washington, DC, 2003) (see www.nap.edu/books/0309088593/html/).
4. P. H. Westfall, S. S. Young, *Resampling-Based Multiple Testing* (Wiley, New York, 1993).
5. C. M. Velicer *et al.*, *JAMA* **29**, 827 (2004).

Response

I THINK MOST WOULD DISAGREE WITH YOUNG'S characterization of the unpublished clinical trials problem as "minor mischief." His larger point, about epidemiological trials, is well taken—but forcing these out of hiding will be administratively difficult. Epidemiological studies undertaken in connection with some regulatory procedures or a court action can be subject to required disclosure. Many and probably most, however, are undertaken in pursuit of a problem that interests the investigator. It might be useful, as Young suggests, to have all that stuff out there. Alas, we don't have a process.

DONALD KENNEDY

Oxide Surface Films on Metal Crystals

IN THEIR REPORT "SAMPLE DIMENSIONS influence strength and crystal plasticity" (13 Aug., p. 986), M. D. Uchic *et al.* demonstrate that the dimensions of micrometer-sized metal crystals have dramatic effects on the properties of the crystals. However, they do not mention a critical factor; namely, that oxide films were present on their specimens. All metals oxidize. Therefore, except for short times in ultra-high vacua, they have oxide (or hydroxide) surface films. These films have negligible effects on the behaviors of most macroscopic metals. As sizes decrease, however, they become increasingly important. The necessity for shearing through these films may well account for the increase in jaggedness of the stress-strain curves observed by the authors with decreasing size. Also, this interpretation is consistent with the lack of a "size effect" observed for the nickel-based superalloy, as well as the large glide offsets.

JOHN J. GILMAN

Department of Materials Science and Engineering, University of California, Los Angeles, 6532 Boelter Hall, Los Angeles, CA 90095, USA.

Response

GILMAN POINTS OUT THAT A THIN SURFACE FILM (oxide or hydroxide) on a micrometer-sized crystal could affect the sample flow behavior, especially as the sample size decreases. He raises an important point. However, it is not entirely clear to us how a surface film could contribute significantly to the strength under progressively rising stresses, without also leading to severe strain localization. High-resolution scanning electron micrographs of deformed samples generally show slip traces at many points along the sample gage length. One might expect much more slip localization in these samples if oxide films were acting as strong barriers to deformation.

The "jaggedness" of the stress-strain response that Gilman mentions may be

simply the result of mechanical testing at the micrometer scale. A slip event in a micrometer-sized crystal will be more clearly resolved than if the same event occurred in a millimeter-sized sample.

In addition, Gilman indicates that the lack of a "size effect" in nickel-based superalloys is consistent with the interpretation that surface films are affecting the observed mechanical behavior. Although this is open to interpretation, we have recently learned through additional testing that sample size effects of the type discussed in our Report do occur for a nickel-based superalloy at sample sizes that are 5 μm in diameter and smaller (*1*).

Finally, significant sample size effects have recently been observed in testing of 1-μm-diameter single crystals of pure gold (*2*), using the same fabrication and testing methodology. The magnitude of the size-dependent strengthening is similar to the results for pure nickel. However, gold does not form a native oxide. Clearly, more research is needed to positively establish the connection between sample size effects to either changes in deformation mechanisms or extrinsic effects such as surface case hardening.

MICHAEL D. UCHIC,[1] DENNIS M. DIMIDUK,[1]
JEFFREY N. FLORANDO,[2] WILLIAM D. NIX[3]

[1]Air Force Research Laboratory, Materials & Manufacturing Directorate, Wright-Patterson Air Force Base, OH 45433–7817, USA. [2]Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [3]Department of Materials Science and Engineering, Stanford University, Stanford, CA 94305–2205, USA.

References
1. M. D. Uchic, D. M. Dimiduk, in preparation.
2. J. R. Greer, W. C. Oliver, W. D. Nix, presentation at the 2004 Fall Minerals, Metal, and Materials Society Meeting, New Orleans, LA, 26 to 29 Sept. 2004.

Protein Design: Quo Vadis?

IN THEIR LETTER "PREDICTING PROTEIN structures accurately" (11 June, p. 1597), M. von Grotthuss *et al.* discuss the ab initio–designed Top7 protein ("Design of a novel globular protein fold with atomic-level accuracy," B. Kuhlman *et al.*, Research Articles, 21 Nov., p. 1364). The discussion in this Letter and in the Response by P. Bradley *et al.* (p. 1597) revolves around the difference between predicting a "real" and a designed structure. Obviously, a fold that has never been seen before cannot be predicted by software that compares with known folds in the Protein Structure Database (PDB), nor necessarily the fold of a sequence for which PDB files with exactly that fold are available.

The underlying problem is that evolution worries about long-term survival of the species and not about protein stability or structural perfection. As stability was never nature's goal, it is not difficult for us to make proteins more stable by mutagenesis. And Kuhlman *et al.* have shown that it also is no longer difficult to design a stable globular protein.

Around 1990, the European Molecular Biology Laboratory organized several "ab initio protein design courses" (*1*, *2*). The participants used the same principles as Kuhlman *et al.* These designs suffered from databases that were too small, primitive software, insufficient CPU power, too little time, and lack of experience. Later we sent the sequence of one of the designs to the first CASP modeling competition (*3*, *4*). All predictors found the correct fold with such high confidence levels that they worried that this so-called "mystery protein" might be a joke. We synthesized the mystery protein, and CD and NMR data showed the right amount of alpha-helix. But it surely did not fold properly, which shows that even when software tells you that things are perfect, it still can be very wrong. And that makes Top7 very special.

GERT VRIEND

CMBI, RU, Toernooiveld 1, Nijmegen 6525 ED, Netherlands. E-mail: Vriend@CMBI.RU.NL

References
1. C. Sander, G. Vriend, Eds., *PRODES90 Protein design on computers*, EMBL Biocomputing Technical document 6 (European Molecular Biology, Laboratory, Heidelberg, Germany, 1991).
2. C. Sander *et al.*, *Proteins* **12**, 105 (1992).
3. See http://predictioncenter.llnl.gov/.
4. Fifth Meeting on the Critical Assessment of Techniques for Protein Structure Prediction, E. E. Lattman, Ed., *Proteins* **53** (issue 6, suppl.), 333 (2003).

CORRECTIONS AND CLARIFICATIONS:

News of the Week: "Researchers build quantum info bank by writing on the clouds" by C. Seife (22 Oct., p. 593). The name of Klaus Mølmer, of the University of Aarhus in Denmark, was misspelled.

2004 Presidential Forum: (1 Oct., p. 46). Some readers have pointed out that an error in the wording of a question on creationism could be interpreted to suggest that *Science* believes intelligent design and creationism are scientific critiques of evolutionary theory. That is not the case. The question should have read: "Should 'intelligent design' or other *un*scientific critiques of evolutionary theory be taught in public schools?" We regret the error.

News Focus: "Finding reactions in a haystack: try 'em all, see what works" by R. F. Service (10 Sept., p. 1558). In the final panel of the diagram, the red dot should have indicated reactants A1 and B4 instead of A3 and B2.

Published by AAAS

HISTORY OF SCIENCE

Solid Tools for Visualizing Science

Julie K. Brown

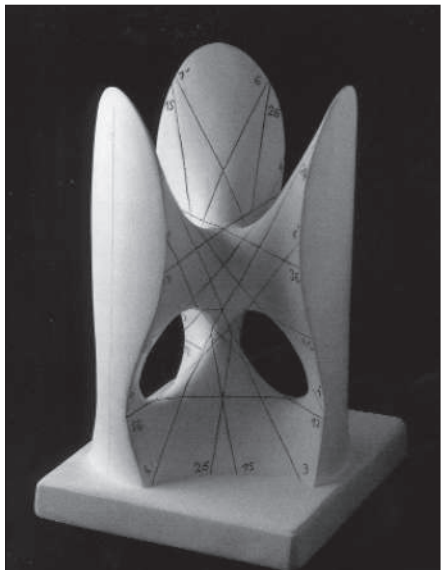
The general public has long been fascinated with three-dimensional models, for what they represent as well as for their sheer beauty and ingenious construction. However, until recently such objects have received little notice from historians of science and medicine. Now, with increasing attention being paid to the visual and material culture of science, scholars finally have begun to examine three-dimensional models for their part in "knowledge production." *Models: The Third Dimension of Science*, a collection of fascinating essays that grew out of a 1998 symposium in London hosted by the Wellcome Institute for the History of Medicine and the Science Museum, stands as the most important contribution to this topic to date.

Models
The Third Dimension of Science
Soraya de Chadarevian and Nick Hopwood, Eds.
Stanford University Press, Stanford, CA, 2004. 482 pp. $65, £45.95. ISBN 0-8047-3971-4. Paper, $24.95, £17.50. ISBN 0-8047-3972-2. Writing Science.

In bringing together the work of 17 international historians of science who have previously written on their respective topics, the volume offers an opportunity to view both the continuity and the diversity in the uses of three-dimensional models. This finely focused theme and the historical specificity of each author's contribution sustain and propel the reader through this rich, provocative collection. Unraveling the how, what, where, and why of three-dimensional models that represent aspects of archaeology, medicine, chemistry, biology, mathematics, and economics does not necessarily produce a consensus. But it does show, as editors Soraya de Chadarevian and Nick Hopwood point out, that models operate on many levels within social and cultural "networks of production and communication." Although the chapters are chronologically arranged—with the focal examples drawn from the 18th century through the mid-20th century—the individual contributors move well beyond fixed historical frameworks. Two final commentaries provide additional context to the topic: James Griesemer offers a philosophical perspective on models and Ludmilla Jordanova discusses the relation of models to visual culture in general.

One key point of the volume concerns the differing yet often overlapping functions of three-dimensional models: as research tools ("epistemic objects") in the making of knowledge, as mediators of information for communication, and as visual objects for public display. Herbert Mehrtens points out that physical models were research tools for some German mathematicians such as Felix Klein but only during a short window of time, in the 1890s. However, their use, Mehrtens notes, as teaching tools, indicators of professional status, and aesthetic forms continued well into the 20th century. In molecular biology the story is more intricate and layered, as Soraya de Chadarevian makes clear. Three-dimensional modeling was already an integral tool and an established research practice for crystallographers when, in the 1950s, molecular biologists adapted it for their studies of proteins. For publication purposes, however, these models had to be transposed back into two dimensions by specialized illustrators. The life of these three-dimensional research models was further extended as they appeared in displays at lectures and exhibitions, entered museum collections, and even were featured in television programs, each time entering a new cultural context where their meaning was redefined.

Rationalizing three-dimensional models as educational tools has been a recurrent theme in their history, because of the fluctuating values attributed to tactile and spatial experiences in the learning process. As James Secord points out, it was the rational education theories of Johann Pestalozzi combined with the Victorian commercial capitalism of the mid-19th century that were responsible for the famous Crystal Palace (Sydenham) display of prehistoric animal models. However, Secord finds that the enterprise, "the apotheosis of a short-lived conjunction," ultimately failed in its intention of bringing the sciences together in a display format popular with the public. Documenting a more complex version of this story, Lynn Nyhart discusses the contrasting pedagogical approaches to three-dimensional models and dioramas taken by two German natural history museums. The use of these "biological groups," she notes, reflected the escalating cultural tensions between the "real" and the "artificial" as well as the growing rift between scientific authority and the demands of popular entertainment in the early 20th century.



Mathematical model, 1986. The four plaster models (here, Clebsch's diagonal surface) reissued to accompany (*2*) served as objects of beauty rather than research tools.

The processes of crafting and constructing three-dimensional models provide a focus for several authors. Nick Hopwood discusses the wax embryological models created by Adolf and Friedrich Ziegler, drawing on his extensive documentation of their studio (*1*). Besides offering excellent descriptions of the process for making wax anatomical models and Felice Fontana's 18th-century collection in Florence, Renato Mazzolini also redirects attention to Fontana's less-known but exceptional "demountable models." These wooden models, which Fontana developed late in life, epitomized his growing belief that learning by touch and assembling physical components superceded the purely visual experience. Thomas Schnalke, who has written widely on 19th-century moulages (clinical "pictures of diseases in wax"), gives the clearest and most layered discussion of model-making as a socio-cultural process. More than a modeling technique, moulage was a process of interaction between the "mouleurs," skilled in taking direct molds of diseased surfaces; the patients, whose participation and presence "never disappeared"; and the supervising physician, for whom the moulage served as a clinical record that enhanced professional status.

The physical disappearance of many of the examples discussed in the volume highlights the intrinsic fragility of these objects as well as the tenuousness of our connection to and understanding of the knowledge that produced them. *Models* reconfirms the value of three-dimensional models and the

The reviewer is the author of *Making Culture Visible*.
E-mail: jkbrown@aol.com

Published by AAAS

continued need to understand how the visual process is integral to both the creation and communication of science.

References

1. N. Hopwood, *Embryos in Wax. Models from the Ziegler Studio* (Whipple Museum of the History of Science, Cambridge, 2002); reviewed by J. Uglow, *Science* **297**, 1651 (2002).
2. G. Fischer, Ed., *Mathematische Modelle / Mathematical Models* (Vieweg, Braunschweig, 1986).

HISTORY OF SCIENCE

One Path Through Los Alamos

Gregg Herken

Physicist Robert Oppenheimer would have been 100 years old last April—had he not died prematurely at age 62, a victim of throat cancer. Any centenary is traditionally a celebratory occasion, and "Oppie" has been the subject of several previous near-hagiographies. David Cassidy, a Hofstra University historian of 20th-century physics and the author of a well-received biography of Werner Heisenberg, has written a book that neither praises Oppenheimer nor buries his reputation but, rather, puts some tarnish upon the icon.

J. Robert Oppenheimer and the American Century
by David C. Cassidy
Pi Press, New York, 2004. 480 pp. $27.95, C$39.95. ISBN 0-13-147996-2.

As Cassidy notes, he meant *J. Robert Oppenheimer and the American Century* to be in keeping with the recent trend in biography, which shifts the focus from the individual toward his or her times. To that end, Cassidy borrows the theme of "scientific militarism" from sociology and the concept of the "American Century" from *Life* publisher Henry Luce. But neither thread really becomes the book's leitmotif, and both are only spottily referenced in the text. This may be because Oppenheimer is a difficult subject for such an approach—he was, arguably, unique (certainly his like had not been seen before, nor has it since)—and also because the times turned out to be as complex as the man.

Although more a scientific biography than a cradle-to-grave life, the book begins with the young Oppenheimer as a way of illuminating its subject's seemingly inexplicable behavior as an adult. It has long been debated whether and how Oppenheimer's Jewishness and his education at Manhattan's Ethical Culture School affected him in later life, but Cassidy makes a persuasive case that each had an effect greater than previous biographers—or Oppie himself—realized. As physicist Isidor Rabi, Oppenheimer's lifelong friend, was fond of observing, being Jewish is a fate, not a faith, and, being fate, is inescapable. Oppenheimer's politicalization in the late 1930s proceeded, in part, from his sympathy for the plight of the Jews in Europe. Similarly, Oppie would eventually find in Communism the "City of Light" that graduates of Felix Adler's secular academy sought; the school's early motto was "To give the best to the poorest."

Indeed, a speech that Adler gave during the First World War would later seem hauntingly prophetic of Oppenheimer's own situation. "The time will come," Adler warned, "when the scientist will be considered and will consider himself a disgrace to the human race who prostitutes his knowledge of Nature's forces for the destruction of his fellow men." Adler's admonition was the result of seeing scientists blindly follow soldiers into the great war, wearing lab coats rather than uniforms, but marching in the same serried ranks. Oppenheimer was only twelve when Adler spoke, but the precocious Oppie's own words—especially the poetry he wrote while an undergraduate at Harvard in the 1920s—suggests that he shared Adler's view. Yet, a generation later, Oppenheimer himself would be at the head of another such column, involved in an even more terrible war, fought with still more terrible weapons.

This irony is what gives Oppenheimer's life its magnetic appeal to biographers, and it provides the narrative push that propels Cassidy's story. Along the way, however, the author journeys down numerous side roads—including education reform movements in the United States, comparisons of science in Europe and America, and U.S. immigration patterns in the early 20th century—that, while often interesting and relevant to the tale, dilute its impact.

Historians of science and practicing scientists will debate whether Cassidy has accurately assessed Oppenheimer's contribution to physics. As have others, Cassidy argues that had Oppie been better focused, and had the war not intervened, his work on stellar evolution might have led him to the discovery of so-called black holes and possibly even to a Nobel Prize. But Cassidy does not grant Oppenheimer the easy deference that prior biographers have. In his view, Oppie was not a leading figure in quantum physics, but simply adept at "mopping up" after the first wave had passed. Oppenheimer even fares poorly in comparison with Heisenberg. Thus, by 1929—when Oppie was being lured by both Caltech and Berkeley—"he was becoming known," writes Cassidy, "not only for his ideas but also for his carelessness in calculations." Elsewhere, Cassidy sees Oppenheimer as "the enabler, but not the prime mover, of fundamental innovation."

Despite his focus more on Oppenheimer as scientist than as political figure, Cassidy makes an effort to include some of the recent evidence and controversies that have surfaced regarding Oppie's protracted dalliance with Communism. But, by including all possible theories and embracing none, the result may be more confusing than enlightening to the reader.

Cassidy's research is generally both thorough and careful, but he too quickly accepts some claims without corroborating evidence. The author blithely passes along the gossip that Oppie had sexual designs on not only Linus Pauling's wife but on Pauling himself. The source for this rather remarkable claim is, evidently, speculation by Pauling's biographer. Cassidy also adds some careless errors of his own to the book: Oppie's brother Frank was anything but "a very successful rancher," and Oppenheimer's friend Haakon Chevalier, to the best of my knowledge, was never "an adviser to DeGaulle."

While hardly demonology, neither is *Oppenheimer and the American Century* a very flattering portrait. It is, rather, a rebuttal to the kind of mid-century hero-worship that tended to portray Oppie as the Galileo of Cold War America. But it is also a portrait that seems, on balance, a bit too harshly drawn. One reason for Cassidy's disillusionment with Oppenheimer may be the disappointment that follows unmet expectations. Ultimately, what Oppenheimer did in putting his science at the service of the state made him not so very different from Edward Teller, Ernest Lawrence, or any other scientist at Los Alamos. The problem is that, with Oppenheimer, somehow, one expected more. As Rabi also observed, puckishly: "People like to think that Oppenheimer built only *good* bombs."

CREDIT: J. ROBERT OPPENHEIMER MEMORIAL COMMITTEE

Partial vindication. Oppenheimer received the Fermi Award, the U.S. Atomic Energy Commission's highest honor, in December 1963.

The reviewer, the author of *Brotherhood of the Bomb*, is a historian at the University of California, Merced, Post Office Box 2039, Merced, CA 95344, USA. E-mail: gherken@ucmerced.edu

Published by AAAS

MOLECULAR BIOLOGY

NIH Molecular Libraries Initiative

Christopher P. Austin,[1*] Linda S. Brady,[2] Thomas R. Insel,[2] and Francis S. Collins[1]

The purpose of the Molecular Libraries Initiative (MLI) component of the NIH Roadmap for Medical Research (*1*, *2*) is to expand the availability, flexibility, and use of small-molecule chemical probes for basic research. Because this initiative is particularly novel and far-reaching, it has been the subject of considerable discussion (*3*–*5*), and sometimes misinterpretation (*6*), in the research community.

Two imperatives motivated the development of the MLI. The first, related to NIH's mission in basic biomedical research, was the need for fundamentally new approaches to determine function and therapeutic potential for all genes in the newly sequenced human genome. The second, related to NIH's mission to improve public health, was the need to accelerate the translation of basic research discoveries into new therapeutics.

Expense (particularly large capital costs), expertise, and cultural divides between public and private sectors have historically kept discovery and optimization of small molecules largely restricted to pharmaceutical and biotechnology companies. Dissemination of small-molecule research tools and technologies into the public sector via the MLI is timely for several reasons.

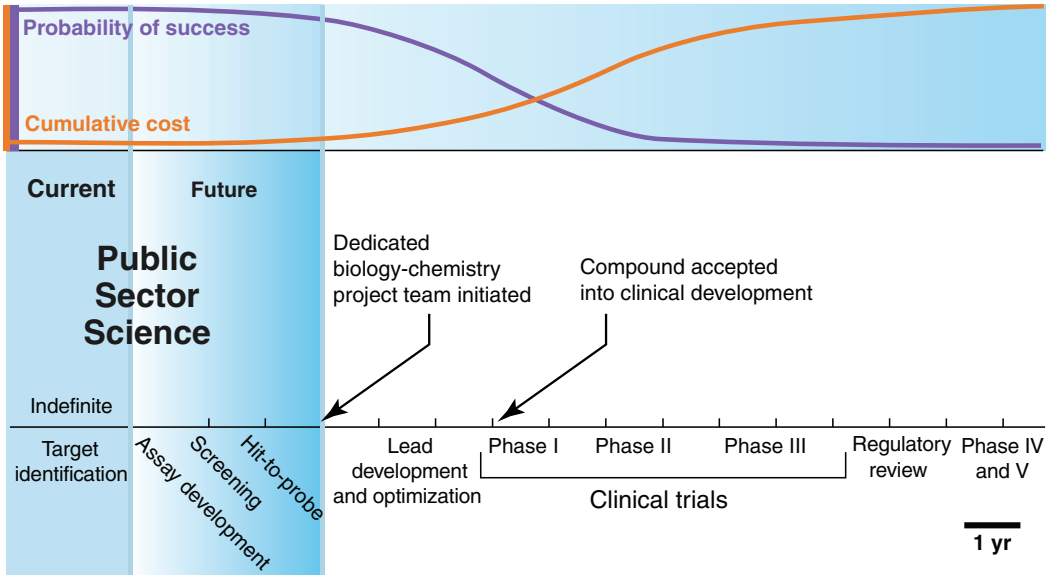
First, sequencing of the human genome has provided an abundance of new targets for study, and small-molecule research tools will accelerate the translation of genome sequence into biological and therapeutic insights (*7*). Small molecules are complementary to nucleic acid–based translational tools such as knockout mice and siRNAs, in that they target the protein gene product rather than the gene locus or mRNA, have virtually limitless structural diversity, can affect particular target functions for defined periods in isolated proteins, cells, or organisms, and can serve as either agonists or antagonists. The characteristics that make this class of molecule useful as drugs—their potential for selectivity, cell permeability, and subtle reversible modulation of important physiological functions—also make them good research tools for dissecting the functions of novel genes, pathways, and cells.

The human genome encodes 20,000 to 25,000 genes (*8*) and perhaps a million proteins, of which only ~500 are targeted by currently available small molecules (*9*). For the most part, pharmaceutical and biotechnology companies prefer to focus on the "druggable genome" thought to be more amenable to drug development (*10*). The majority of the genome that is currently considered "undruggable" (i.e., unmanipulable by small molecules) is therefore a major focus of the MLI.

Large libraries of small molecules have traditionally been unavailable to academic researchers, but with the advent of combinatorial chemistry and commercial suppliers of high-quality compound libraries, small molecules can now be obtained on a large scale. At the same time, advances in robotics and informatics have made screening and analysis of such large compound libraries possible. Up to a million compounds can now be screened against a target in a single day, three orders of magnitude greater than was possible only a decade ago. Together, these developments make a public-sector small-molecule screening and chemistry initiative such as the MLI possible.

The MLI was developed over the course of 9 months through consultations with representatives of multiple NIH institutes, and external consultants from the public and private sectors. The MLI research agenda has three components focused on screening, cheminformatics, and technology development, and is being carried out via NIH grant and contract mechanisms (*11*).

The Molecular Libraries Screening Center Network (MLSCN) will be a consortium of five or six high-throughput screening (HTS) centers that will screen assays submitted by the research community on a large number of compounds (>100,000) maintained in a central compound repository, and perform optimization chemistry required to produce in vitro chemical probes of the targets or phenotypes studied in the assays (*12*). All results will be placed into a new public database (PubChem, see below), and probe compounds will be made available without encumbrance to all researchers, in public and private sectors, for their use in studying bi-



Interface of the MLI and drug development.

[1]National Human Genome Research Institute and [2]National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892, USA. *Author for correspondence. E-mail: austinc@mail.nih.gov

ology and disease. The first of the MLSCN screening centers (the NIH Chemical Genomics Center) has been established within the NIH intramural program. It will begin full-scale screening in early 2005 (*13*). The other MLSCN centers will be funded via extramural grants; applications to this NIH Roadmap Request for Applications (*12*) are under review and will be awarded in the late spring of 2005. The contract for the Molecular Libraries Small Molecule Repository, which will house the screening collection, was recently awarded (*14*), and the composition of the compound collection is being determined by a distinguished panel of chemists from the public and private sectors.

A comprehensive database of chemical structures and their activities, PubChem (*15*), has been developed by the National Center for Biotechnology Information at the National Library of Medicine/NIH. PubChem links small-molecule information to GenBank, MEDLINE, and the other Entrez databases, and will serve as a public portal for MLSCN screening results and chemistry data. New algorithms and tools for computational chemistry, virtual screening, and other research aspects of cheminformatics, will be funded via new grants (*11*).

The ultimate goal of the MLI is to develop small-molecule modulators of thousands of cellular targets. To succeed, the MLI will be developing technology in four critical areas (*11*):

(i) **Chemical diversity**: production of pilot-scale compound libraries in novel areas of chemical space (*16*), and methods development for natural product isolation, characterization, and chemistry. Compounds will be placed into the Small Molecule Repository and screened by the MLSCN centers.

(ii) **Assay diversity**: development of innovative high-throughput assays for novel proteins, cellular phenotypes, biological functions, and disease mechanisms.

(iii) **Instrumentation**: development of new technologies to allow HTS of novel assay formats and to increase throughput and accuracy of current screening technologies (e.g., methods for highly parallel noncompetitive detection of target binding, lab-on-chip technologies enabling complex multistep assays).

(iv) **Predictive ADME/Toxicology:** development of data sets and analysis algorithms for improved prediction of ADME (absorption, distribution, metabolism, and excretion) and toxicity properties of small molecules.

Although the MLI will utilize tools and technologies found in biopharmaceutical companies, the MLI has quite distinct goals and deliverables from those of the private sector. For example, the need to develop compounds for human use as quickly as possible drives biopharmaceutical companies to screen only "druglike" compounds having potential for intellectual property novelty, on assays for a relatively limited group of targets (e.g., GPCRs, kinases, nuclear hormone receptors), and to not disclose the results of screens, keeping them as trade secrets or intellectual property. In contrast, the types of compounds synthesized and screened by the MLI will be broader, including metabolic intermediates, a range of natural products, and agents with potential in vivo toxicity. Types of assays will be broader also, including, for example, protein-protein interactions, splicing events, and diverse cellular and even organismal phenotypes. Via PubChem and the compound repository, all data and chemical probes will be available to the entire research community without encumbrance. The MLI will thereby enable study of the majority of biological and chemical space that is currently unexplored.

The small-molecule research tools produced by the MLI should accelerate validation of new drug targets (*17*) and thereby enable new drug development, but clinically useful compounds cannot be expected from the MLI itself. Drug development is a complex, time-consuming, and expensive process (*18, 19*), only the first steps of which will be performed by the MLI (see the figure on page 1138). The "probe compounds" produced by the MLI will have potency and aqueous solubility adequate for in vitro applications, but chemical modifications will generally be needed to confer the selectivity, pharmacokinetic, and metabolic properties required for *in vivo* use. From the probe stage, 10 to 12 chemist-years are commonly required to develop a "lead compound" with minimal pharmaceutical properties, and an additional 20 to 30 chemist-years to produce a "clinical candidate" compound appropriate for testing in humans. During this time, >3000 different compounds based on the initial probes are typically synthesized and tested. Even after this investment in chemical optimization, >90% of clinical candidates fail in human testing (*19*).

Examination of the cost and expected success rate at each stage of drug development demonstrates that the assay development, HTS, and hit-to-probe chemistry steps to be performed by the MLI are inexpensive and straightforward compared with the later stages of drug development (see the figure on page 1138). This should allow the MLI to produce probes for a broad range of targets with its relatively limited budget of ~$100 million per year (compared with >$30 billion per year spent by pharmaceutical and biotechnology companies on drug development). We estimate that the probes produced by the MLI will entail only 2% of the cost, and 5% of the time, required to develop a novel drug (*19*). In exceptional circumstances, the NIH itself might attempt the subsequent steps of drug development, particularly if a small-molecule probe shows special potential for development into a drug for an orphan indication or a disease occurring primarily in the developing world, where there is unlikely to be commercial interest. In this case, NIH Roadmap mechanisms such as the Translational Research Core Services Program and the Regional Translational Research Centers (*20*) could be used. But for the most part, we expect that chemists in the public and private sectors will use the MLI probes as proof-of-concept compounds and as starting points to produce a variety of chemical analogs with improved properties. For this reason, intellectual property claims on the MLI probes will be strongly discouraged, as such claims would prevent the widest use of these tools and is contrary to the "community resource" nature of this initiative (*21*).

The MLI is a bold initiative to catalyze science in the genome era. By providing a new path for discovery, this program aims to accelerate science and its translation into benefits for the health of the public.

References and Notes

1. E. Zerhouni, *Science* **302**, 63 (2003).
2. http://nihroadmap.nih.gov/molecularlibraries/index.asp.
3. J. Kaiser, *Science* **304**, 1728 (2004).
4. D. Janssen, *Bio-IT World* **3**, 1 (2004).
5. F. S. Collins, *Nature Rev. Drug Discov.* **3**, 640 (2004).
6. J. Couzin, *Science* **302**, 218 (2003).
7. C. P. Austin, *Curr. Opin. Chem. Biol.* **7**, 511 (2003).
8. International Human Genome Sequencing Consortium, *Nature* **431**, 931 (2004).
9. J. Drews, *Science* **287**, 1960 (2000).
10. A. L. Hopkins, C. R. Groom, *Nature Rev. Drug Discov.* **1**, 727 (2002).
11. http://nihroadmap.nih.gov/molecularlibraries/grants.asp.
12. http://grants1.nih.gov/grants/guide/rfa-files/RFA-RM-04-017.html.
13. www.genome.gov/12512295
14. http://pubs.acs.org/cen/news/8235/8235notw5.html.
15. http://pubchem.ncbi.nlm.nih.gov/.
16. N. P. Savchuk, K. V. Balakin, S. E. Tkachenko, *Curr. Opin. Chem. Biol.* **8**, 412-417 (2004).
17. M. A. Lindsay, *Nature Rev. Drug Discov.* **2**, 831-834 (2003).
18. K. H. Bleicher, H. J. Bohm, K. Muller, A. I. Alanine, *Nature Rev. Drug Discov.* 2, 369-78 (2003).
19. J. A. DiMasi, R. W. Hansen, H. G. Grabowski, *J. Health Econ*. **22**, 151 (2003).
20. http://nihroadmap.nih.gov/clinicalresearch/overview-translational.asp
21. http://grants1.nih.gov/grants/guide/notice-files/NOT-RM-04-014.html.
22. The Molecular Libraries Roadmap is a trans-NIH initiative conceptualized and implemented with the participation of all Institutes and Centers; see http://nihroadmap.nih.gov/molecularlibraries/members.asp for the members of the Molecular Libraries Project Team. The authors are leaders of the initiative.

ECOLOGY

How Extinction Patterns Affect Ecosystems

David Raffaelli

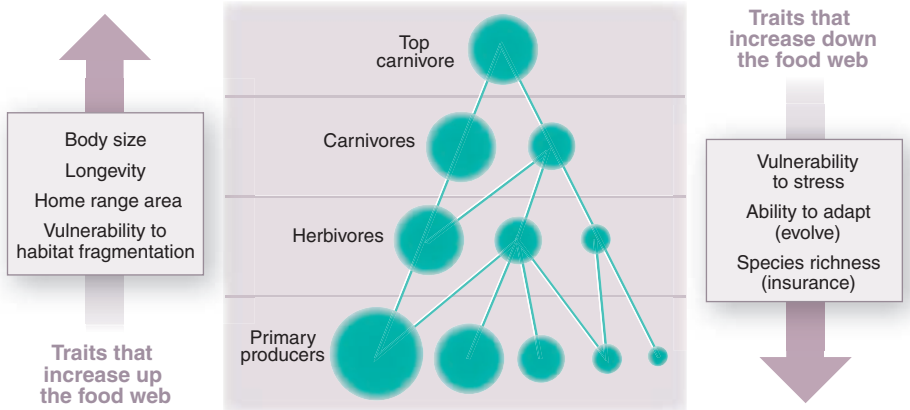
We live in a world where the accelerated extinctions of species and changes in biodiversity are no longer disputed issues. Much effort has gone into quantifying biodiversity loss rates for particular animal and plant groups (*1*). Less clear, however, is the impact of such losses on ecosystems, especially when many different kinds of species of plants and animals are lost simultaneously (*2*). Yet policy-makers urgently need guidance on the effects of multispecies losses if they are to plan for and advise on the societal consequences of biodiversity changes. The ecological research community has been highly active in attempting to provide such guidance (*3*, *4*), but many challenges remain. Foremost among these is that most real extinction events are nonrandom with respect to species identity—some species are more likely to go extinct than others—whereas research studies often assume that extinctions are random. Two papers in this issue, by Solan *et al*. on page 1177 (*5*) and Zavaleta and Hulvey on page 1175 (*6*), reporting on work in two very different types of ecosystem, reveal that the impact of nonrandom species extinctions on ecosystems is markedly different from that predicted by scenarios where extinctions are random. These studies bring us a step nearer to understanding the impact of nonrandom species losses on ecosystems and should help to provide policy-makers with a firmer basis for decision-making.

The two studies examine very different habitats (marine versus terrestrial), each with different kinds of organisms (sea-bed invertebrates versus grassland plants), different ecosystem processes (sediment biogeochemistry versus resistance to invasion by exotic species), and different types of experimental approaches (data analysis and modeling versus controlled experimentation). So it is all the more interesting, for scientists and policy-makers alike, that both papers arrive at the same conclusion: Nonrandom extinction events have impacts on ecosystems that are quite different from those predicted by scenarios that assume species extinctions occur at random.

Solan *et al*. (*5*) combine into an elegant model an extremely well-documented data set of invertebrate communities in marine sediments off the coast of Galway, Ireland (*7*). This fusion, facilitated by the BIOMERGE initiative (*8*), enables the authors to predict what will happen to the cumulative effects of the small-scale sediment disturbances (bioturbation) caused by the movement, feeding, and respiration activities of all 139 species of clams, worms, sea urchins, brittle stars, and shrimps present in this system if species are lost through impacts such as overfishing, habitat destruction, and pollution. They scored each species for its body size, mobility, and mode of sediment mixing to calculate an index of bioturbation potential for different species combinations and for different degrees of species richness. In their model, either extinction scenarios could be random or losses could be ordered with respect to the sensitivity of species to environmental stress, body size, and abundance, traits that in turn reflect different kinds of impact (see the figure).

Not surprisingly, both types of scenario reduced sediment bioturbation, but the order in which species disappeared had a greater overall impact on the ecosystem. One particular species of brittlestar (a relative of the starfish) was clearly a keystone species. Due to its high abundance, large body size, and high mobility, the loss of this brittlestar species affected bioturbation disproportionately. When this species disappeared, the ecosystem changed dramatically. In contrast, the loss of other species that were small or rare had only minor effects on the ecosystem. Thus, rare or small-bodied species are unable to compensate for declines in ecosystem processes when larger, more abundant species are lost, contrary to earlier suggestions. Clearly, the effects of biodiversity loss on ecosystem integrity will depend largely on the order in which species are lost, which in turn is determined by the susceptibilities of ecosystems to different types of stresses (overfishing, habitat fragmentation, and pollution).

In contrast to the modeling strategy used by Solan *et al*., Zavaleta and Hulvey



Extinctions are rarely random in the real world. In the food web shown, the loss of biodiversity has a different impact depending on the trophic level, because the traits of these species that make them vulnerable to different impacts covary both between and within trophic levels. Thus, the body size (represented by the size of the circles) of top carnivore species tends to be larger than that of species at lower trophic levels (predators are usually larger than their prey). Similarly, there is a range of body sizes (and other traits) within each trophic level, and there are also more species represented at lower trophic levels. Traits such as body size and its covariates such as home range and tolerance to stress, together with differences in species richness between trophic levels, will determine the impact on ecosystems of different biodiversity loss scenarios. For instance, top predators with their large body size, low abundance, and large home range requirements are particularly vulnerable to habitat fragmentation/destruction, but may be less susceptible to contaminant stress, which affects smaller species disproportionately. In contrast, because of the higher species richness at lower trophic levels, there is more "insurance" against the effects of species losses, and these species may also have a greater capacity for adaptive change due to their shorter life-spans and faster turnover rates. With respect to species traits, extinction is unlikely to occur randomly (*5*, *6*).

The author is in the Environment Department, University of York, York YO10 5DD, UK. E-mail: dr3@york.ac.uk

(*6*) adopted a highly controlled experimental approach. They used a relatively simple (21 species) assemblage of California grassland plants, which they arranged in treatment plots to provide a range of species richness. Half of the plots were challenged by an exotic invader, the yellow star thistle, to assess the resistance of the different communities to invasion (an ecosystem characteristic). Their experimental design constitutes a major advance over previous experiments on grassland systems in that the species richness treatments that they assembled represent plant communities at different stages of biodiversity loss and reflect nonrandom extinctions. The authors achieved this by applying nested subset analysis to a 4-year data set of spatial variation in plant communities in order to reveal those changes in richness that occur in a consistent, nested order. These changes were then reflected in the setup of the experimental plots.

The results were dramatic: Resistance to invasion by the thistle was tightly correlated with the species richness of the invaded community. These findings are completely at odds with the results of similar experiments in which species richness treatments were assembled randomly. The authors conclude that the consequences of species loss may be greater than that predicted on the basis of studies in which extinctions are assumed to occur randomly.

Both the Solan *et al*. and Zavaleta and Hulvey studies were motivated by the need to inject greater realism into analyses of how biodiversity affects ecosystems. Many of the high-profile studies in this area are concerned with the measurement of a response variable (ecosystem process) under conditions of different numbers of species assembled randomly and usually comprising a single trophic level. Although such research provides a bivariate plot of ecosystem integrity versus species richness, it is uncertain whether such relationships reflect what happens when species are lost from ecosystems. Controlled removal of species from species-rich assemblages to produce a lower level of richness would go a long way toward answering this question. Species removal, however, is fraught with practical obstacles and difficulties over interpretation of results, nothwithstanding the issue of random versus nonrandom extinctions. These two studies illustrate how it is possible to use alternative approaches to tackle this aspect of the biodiversity-ecosystem dynamic more realistically, but much remains to be done. Zavaleta and Hulvey focus on a single trophic level (plants). Solan *et al*. concentrate on a community that comprises different trophic types, but this community does not capture the larger food web that includes fish and other consumers. Different ecological impacts will affect biodiversity differently at various trophic levels (see the figure). The major challenge facing ecologists is to be able to advise policymakers about the effects of such combined losses on ecosystems throughout a food web. These two papers take us an important step nearer to achieving that goal.

References and Notes

1. See www.royalsoc.ac.uk/events.
2. See www.millenniumassessment.org/en/index.aspx.
3. M. Loreau *et al*., *Science* **294**, 804 (2001).
4. M. Loreau, S. Naeem, P. Inchausti, *Biodiversity and Ecosystem Functioning* (Oxford Univ. Press, Oxford, 2002).
5. M. Solan *et al*., *Science* **306**, 1177 (2004).
6. E. S. Zavaleta, K. B. Hulvey, *Science* **306**, 1175 (2004).
7. J. L. Ruesink, D. S. Srivastava, *Oikos* **93**, 221 (2001).
8. See www.columbia.edu/cu/biomerge.

PHYSICS

Superfluid Helium-3 Has a Metallic Partner

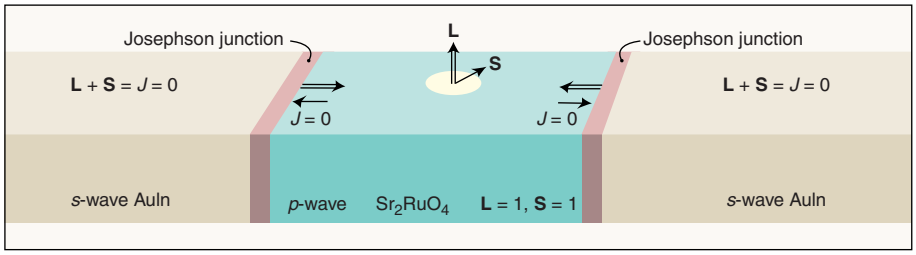
Maurice Rice

After the unexpected discovery of high-temperature superconductors, it took almost a decade and a new class of experiments (*1*) to establish that their electronic nature was fundamentally different from that of conventional, low-temperature superconductors. Now, Nelson *et al.*, on page 1151 of this issue, (*2*) have performed similar experiments on strontium ruthenate (Sr_2RuO_4), a metallic superconductor that does not conform to either of the two types. In an experimental tour de force, the authors confirm that this ruthenate metal is the long-sought metallic analog of superfluid helium-3 (^{3}He).

In a superconductor, electron pairs move through the material without encountering any electrical resistance. The electrons move both as a pair and relative to each other. According to the original Bardeen-Cooper-Schrieffer theory (*3*), the pairing is caused by interactions between the electrons and the crystal lattice. The relative motion of the paired electrons can be described by an *s*-wave structure, corresponding to an angular momentum $l = 0$. Not long after this theory was developed, Kohn and Luttinger speculated that superconductors in which the electrons have finite angular momentum ($l = 1$ or higher) could also occur (*4*). It was unclear, however, how such an unconventional superconductor could be stabilized and how its unconventional nature could be proved.

Early efforts focused on the first question. It was proposed that pairing could arise from the exchange of spin or from fluctuations in the magnetization rather than from lattice vibrations. This idea received a boost when superfluidity was discovered in the early-1970s in liquid ^{3}He and was shown to be a pair condensate with *p*-wave symmetry ($l = 1$). Just like electrons, the ^{3}He isotope is a fermion (that is, no two ^{3}He atoms can occupy the same state at the same time). The ^{3}He condensate is called a superfluid rather than a superconductor simply because, in contrast to electrons, the helium atoms do not carry a charge.

However, the search for a metallic counterpart to superfluid ^{3}He remained fruitless at first. In the 1980s, superconductivity was observed in metals that did



The definitive experiment. The *p*-wave superconductor Sr_2RuO_4 is joined at either end to a conventional *s*-wave superconductor. The angular momentum vector **L** of the Cooper electron pairs in Sr_2RuO_4 is oriented along the *c* axis of the tetragonal crystal, whereas the spin vector **S** lies in the *ab* plane. At the insulating barriers, **L** rotates in the *ab* plane, and $J = 0$ electron pairs are formed. The formation of these $J = 0$ pairs allows interconversion between the spin-singlet pairs in the *s*-wave superconductor and the spin-triplet pairs in the *p*-wave superconductor.

CREDIT: PRESTON HUEY/*SCIENCE*

The author is at the Institut für Theoretische Physik, ETH Hönggerberg, 8093 Zürich, Switzerland. E-mail: maurice.rice@itp.phys.ethz.ch

not fit the Bardeen-Cooper-Schrieffer mold. First came the "heavy fermion" metals, which contain rare earth or actinide ions with local magnetic moments that are only weakly coupled to the conducting electrons. There are many reasons to believe that superconductivity in these metals is unconventional, but their intrinsic complexity has so far prevented a definitive determination of their electronic nature (*5*). Next came the high-temperature cuprates, in which the relative motion of the paired electrons has a *d*-wave structure ($l = 2$). However, cuprates have many additional anomalies that do not fit into a generalized Bardeen-Cooper-Schrieffer theory (*6*).

A decade ago, Maeno *et al.* (*7*) succeeded in growing perfect crystals of Sr_2RuO_4 and found that they were superconducting at low temperatures. Above the superconducting transition temperature, the material behaves as a standard metal, but there are clear signs that the motion of the electrons is strongly correlated at all temperatures, similar to ^{3}He atoms, which move in this way to avoid each other both in the superfluid and the normal liquid. This observation quickly led to speculation that the superconducting state of Sr_2RuO_4 would also be *p* wave ($l = 1$) (*8*). This type of relative motion has consequences for the spin state of the pairs. Basic quantum mechanics for fermions requires the spins of electrons pairs with odd angular momentum to be parallel—that is, to form spin-triplet pairs, rather than the spin-singlet pairs in conventional and cuprate superconductors. Subsequent studies of the magnetic properties of Sr_2RuO_4 pointed to triplet pairing, and much supporting evidence has accumulated since then (*9*).

The final confirmation requires that the internal *p*-wave structure of the pairs be tested in a quantum interference experiment. The way to do this was outlined by Geshkenbein *et al.* (*10*) almost two decades ago. They proposed to modify the standard superconducting quantum interference device (SQUID), which consists of a ring with weak links in the left and right arms. A weak link (called a Josephson junction) can be made by inserting a narrow insulating barrier. Quantum tunneling of the electron pairs through the barrier is usually maximized when their phases on either side are equal. However, Geshkenbein *et al.* (*10*) pointed out that, in the case of *p*-wave superconductivity, a suitably chosen geometry would add an extra phase shift to the tunneling amplitude of one of the two Josephson junctions. The interference pattern of a SQUID is observed by measuring the dependence of the maximum supercurrent through the device as the total magnetic flux flowing through the ring is varied. The extra phase shift is easily recognized as a displacement of the interference pattern.

In their experiments, Nelson *et al.* (*1*), chose a geometry in which half of the ring is a conventional *s*-wave superconductor, which is joined to opposite ends of a Sr_2RuO_4 sample. As illustrated in the figure, tunneling between *s*- and *p*-wave superconductors requires interconversion between spin-singlet and spin-triplet pairs. The observation of the added phase shift by Nelson *et al.* is the final piece of evidence that confirms the close analogy between the metallic superconductor Sr_2RuO_4 and superfluid ^{3}He.

References

1. C. C. Tsuei, J. R. Kirtley, *Rev. Mod. Phys.* **72**, 969 (2000).
2. K. D. Nelson, Z. Q. Mao, Y. Maeno, Y. Liu, *Science* **306**, 1151 (2004).
3. J. Bardeen, L. N. Cooper, J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
4. W. Kohn, J. J. Luttinger, *Phys. Rev. Lett.* **15**, 524 (1965).
5. P. Thalmeier *et al.*, *Condens. Matter*, in press; preprint available at http://xxx.lanl.gov/abs/cond-mat/0409363.
6. P. W. Anderson *et al.*, *J. Phys. Condens. Matter* **16**, R755 (2004).
7. Y. Maeno *et al.*, *Nature* **372**, 532 (1994).
8. T. M. Rice, M. Sigrist, *J. Phys. Condens. Matter* **7**, L643 (1995).
9. A. P. Mackenzie, Y. Maeno, *Rev. Mod. Phys.* **75**, 657 (2003).
10. V. B. Geshkenbein, A. I. Larkin, A. Barone, *Phys. Rev. B* **36**, 235 (1987).

CREDIT: DEE BERGER / PHOTO RESEARCHERS, INC.

OCEAN SCIENCE

Deep Ocean Overturning—Then and Now

Jess F. Adkins and Claudia Pasquero

The deep ocean contains nearly all the mass, thermal inertia, and carbon in the ocean-atmosphere system. The rate at which it overturns can therefore have a profound effect on climate. Measurements of the radiocarbon (^{14}C) distribution in the deep sea suggest that the modern ocean overturns once every ~1000 years. Radiocarbon is an accurate tracer of this process because it is created in the upper atmosphere by cosmic rays and can only enter the ocean through absorption of CO_2 gas at the ocean surface. When the surface water sinks to the deep ocean, it is isolated from the radiocarbon source, and its radiocarbon content decays with its characteristic half-life of 5730 years. Today, there is more radiocarbon in the deep Atlantic Ocean (as a result of deep-water formation in the North Atlantic) than in the Pacific Ocean. We say that the Pacific is "older" or less well "ventilated" than the Atlantic.

Enhanced online at www.sciencemag.org/cgi/content/full/306/5699/1143

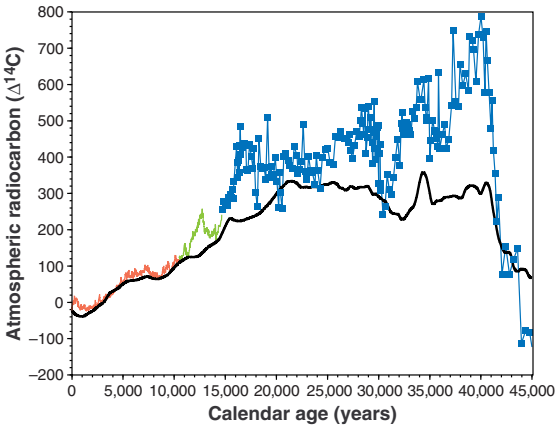
But what about the past ocean, especially at the time of the Last Glacial Maximum, about 20,000 years ago, when large ice sheets covered most of Canada? The calcium carbonate shells of foraminifera (see the first figure) in deep-ocean sediments from this time record the chemistry of the water in which they grew. But here we run into a problem. We cannot use radiocarbon in bottom-dwelling foraminifera as a tracer of water mass age, because their radiocarbon values record the sum of the ages of the sediment and of the water in which they lived. Without an independent chronometer for sediment age, we are stuck with one measurement and two unknowns. However, we can still measure the radiocarbon difference between old samples, as long as we know that they formed at the same time in the past.

A foraminiferan of the kind used in (*1*) to determine the age of mid-depth waters.

This trick is used by Broecker *et al.*, on page 1169 of this issue (*1*), to estimate the radiocarbon age of the equatorial deep Pacific at a depth of 2 km during the Last Glacial Maximum. The authors compare the age difference between surface-dwelling and bottom-dwelling foraminifera both today and at the Last Glacial Maximum. They find that the surface-to-deep age difference was the same, or a little larger, in the past.

This result is somewhat surprising. In a paper published earlier this year in *Science*, Hughen *et al.* came to a different conclusion (*2*) (see the second figure). The authors documented the history of the atmosphere's radiocarbon content over the past 50,000 years. The atmosphere provides an indirect record of the ocean overturning rate, because its radiocarbon budget is very sensitive to how much CO_2 it exchanges with

The authors are in the Department of Geology and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: jess@gps.caltech.edu



Atmospheric radiocarbon content over the past 45,000 years. Data from tree rings [red line (*11*)] and from foraminifera from the Cariaco Basin [green line (*12*) and blue points (*2*)] are compared to modeled production rate variations (*2*) (black line). Deviations between the data and the model reflect changes in the exchange of carbon between active reservoirs relative to today. The two most important processes for the largest deviations are changes in carbonate sedimentation and in the ocean overturning rate. The data suggest that, assuming constant carbonate flux, the glacial ocean overturned more slowly than it does today.

the deep ocean. Land biota and soils also take up radiocarbon, but export to the deep sea is by far the largest sink for modern atmospheric radiocarbon. If this sink were reduced, the atmospheric radiocarbon content would rise, and rise fast. If a new steady state were reached, and the production rate remained constant, the partitioning of radiocarbon between the atmosphere and deep ocean would shift toward higher atmospheric $^{14}C/^{12}C$ ($\Delta^{14}C$) ratios. This is exactly what Hughen *et al.* found for 11,500 to 13,000 years ago (see the second figure, green line) (*2*).

But there are several important other factors to consider. The production rate of radiocarbon is not constant. Radiocarbon production can be tracked through time by measuring the abundance of other radioactive isotopes produced by cosmic rays, such as ^{10}Be and ^{36}Cl, in ice cores (*3*); these isotopes are not affected by carbon sinks. It turns out that radiocarbon production was higher in the past (see the second figure, black line). But the measured $\Delta^{14}C$ at the Last Glacial Maximum was even higher (see the second figure, blue symbols). Hence, the ocean overturning must have been slower. However, Broecker *et al*. show with a simple calculation that the implied increase in the radiocarbon age of the deep ocean is much larger than their measured age difference from the foraminifera.

How can these disparate results be reconciled? One possibility is that the carbon cycle was fundamentally different at the Last Glacial Maximum. Changes in the burial rate of carbonate sediments can affect atmospheric radiocarbon, because these sediments remove radiocarbon from the system. Another option is that the waters below 2-km depth were much older than those measured by Broecker *et al*. There is not much data, but two papers have found very radiocarbon-depleted waters in the deepest parts of the ocean at the Last Glacial Maximum (*4*, *5*). Clearly we need water column profiles of radiocarbon for the Last Glacial Maximum from both the Atlantic and the Pacific.

Finally, it is possible that 19,000 years ago—the age of some of the data used by Broecker *et al*.—the system was not in a steady state. The atmospheric record in the figure shows that large swings in $\Delta^{14}C$ occurred regularly during the last glacial period. Such swings did not occur over the past 10,000 years of relatively constant climate, during which the radiocarbon cycle was in a steady state. Ice core records in Greenland show that no equivalent period occurred during the last glacial period.

Moreover, the small but growing number of deep radiocarbon values from the Last Glacial Maximum (*1*, *4*, *5*) provides insights into what drives the strength of the overturning circulation: Contrary to a widely held belief, the circulation rate is not driven by surface density gradients in the north-south direction. This conclusion agrees with recent theoretical arguments (*6*–*8*). For example, Huang has shown that a modeled increase in the density of high-latitude waters did not directly result in an increased rate of deep-water formation (*8*). In fact, a large meridional surface density gradient induces a strong vertical stratification, which inhibits the return of deep water to the surface and weakens circulation.

Observational data from the paleoclimate record support this theory. The ocean interior was more highly stratified at the Last Glacial Maximum than it is in the modern ocean (*9*), but the circulation was not much stronger, and was possibly slower, than it is today. More data are needed to determine whether the strength of the overturning circulation depends on winds and tidal energy (*7*, *8*, *10*), rather than on surface warming/cooling or evaporation/precipitation budgets at the surface. The growing paleoceanographic data set shows great promise to answer this question.

References

1. W. Broecker *et al.*, *Science* **306**, 1169 (2004).
2. K. Hughen *et al.*, *Science* **303**, 202 (2004).
3. R. Muscheler *et al.*, *Earth Planet. Sci. Lett.* **219**, 325 (2004).
4. N. J. Shackleton *et al.*, *Nature* **335**, 708 (1988).
5. L. Keigwin, *Paleoceanography*, in press.
6. C. Wunsch, R. Ferrari, *Annu. Rev. Fluid Mech.* **36**, 281 (2004).
7. W. Munk, C. Wunsch, *Deep-Sea Res.* **45**, 1977 (1998).
8. R. X. Huang, *J. Phys. Oceanogr.* **29**, 727 (1999).
9. J. F. Adkins, K. McIntyre, D. P. Schrag, *Science* **298**, 1769 (2002).
10. J. R. Toggweiler, B. Samuels, *Deep-Sea Res.* **42**, 477 (1995).
11. M. Stuiver *et al.*, *Radiocarbon* **40**, 1041 (1998).
12. K. Hughen *et al.*, *Science* **290**, 1951 (2000).

EVOLUTION

Genomic Databases and the Tree of Life

Keith A. Crandall and Jennifer E. Buhay

Although we have not yet counted the total number of species on our planet, biologists in the field of systematics are eagerly assembling the Tree of Life (*1*, *2*). The Tree of Life aims to define the phylogenetic relationships of all organisms on Earth. On page 1172 of this issue, Driskell *et al.* (*3*) propose an intriguing computational method for assembling this phylogenetic tree.

These investigators probed the phylogenetic potential of ~300,000 protein sequences sampled from the GenBank and Swiss-Prot genetic databases. From these data, they generated "supermatrices" and then supertrees. Supermatrices are extremely large data sets of amino acid or nucleotide sequences (columns in the matrix) for many different taxa (rows in the matrix). Driskell *et al*. constructed a supermatrix of 185,000 protein sequences for more than 16,000 green plant taxa and one of 120,000 sequences for nearly 7500 metazoan taxa. This compares with a typical systematics study of, on a good day, four to six partial gene sequences for 100 or so taxa. Thus, the potential data enrichment that comes with carefully mining genetic databases is terrific. However, this enrichment comes at a cost. Traditional phylogenetic studies sequence

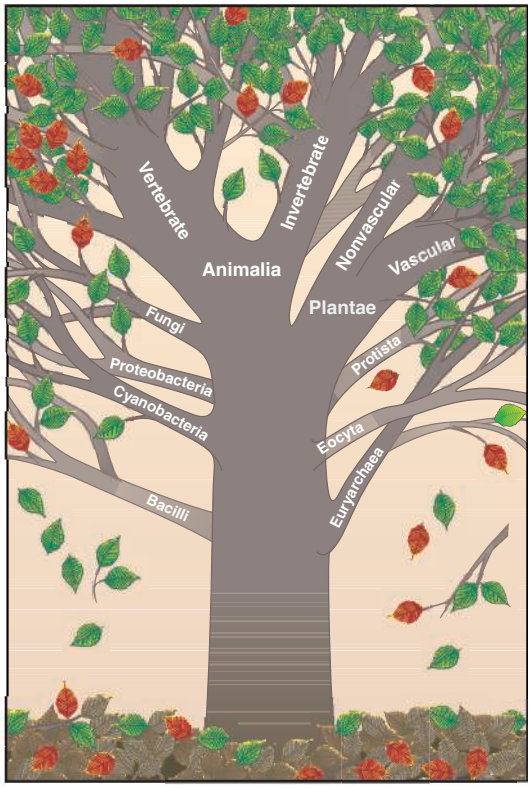
The authors are in the Department of Integrative Biology, and K. A. Crandall is also at the Monte L. Bean Life Science Museum and the Department of Microbiology and Molecular Biology, Brigham Young University, Provo, UT 84602, USA. E-mail: keith_crandall@byu.edu, crayfish@email.byu.edu

the same gene regions for all the taxa of interest while minimizing the overall amount of missing data. With the database supermatrix method, the data overlap is sparse, resulting in many empty cells in the supermatrix, but the total data set is massive.

To solve the problem of sparseness, the authors built a "supertree" (*4*). The supertree approach estimates phylogenies for subsets of data with good overlap, then combines these subtree estimates into a supertree. Driskell and colleagues took individual gene clusters and assembled them into subtrees, and then looked for sufficient taxonomic overlap to allow construction of a supertree. For example, using 254 genes (2777 sequences and 96,584 sites), the authors reduced the green plant supermatrix to 69 taxa from 16,000 taxa, with an average of 40 genes per taxon and 84% missing sequences! This represents one of the largest data sets for phylogeny estimation in terms of total nucleotide information; but it is the sparsest in terms of the percentage of overlapping data. Yet even with such sparseness, the authors are still able to estimate robust phylogenetic relationships that are congruent with those reported using more traditional methods. Computer simulation studies (*5*) recently showed that, contrary to the prevailing view, phylogenetic accuracy depends more on having sufficient characters (such as amino acids) than on whether data are missing. Clearly, building a supertree allows for an abundance of characters even though there are many missing entries in the resulting matrix.

Several questions remain, however, about this strategy. First, the supertree strategy depends fundamentally on our ability to distinguish between orthologous (derived from a speciation event) and paralogous (derived from a duplication event) gene sequences (*6*). The methods to draw this distinction are in their infancy. Little work has been done to compare such methods in terms of their accuracy and their robustness with respect to data that do not fit underlying assumptions (such as neutral evolution). The distinction between the two types of gene sequences typically relies on a well-populated database. Second, supertree approaches themselves are controversial, in part because the methodology results in a degree of disconnect between the underlying genetic data and the final tree produced. Moreover, this strategy has yet to be validated by computer simulation or well-established phylogenetic methods. Third, the supertree approach makes a fundamental assumption: that a bifurcating tree topology represents the genomic evolutionary history of species. This assumption has been called into question because of the reality of genetic exchange across species boundaries through mechanisms such as horizontal gene transfer and hybridization. Depicting genealogical relationships as networks might better represent the true underlying biology (*7*, *8*).

Nonetheless, the ability of Driskell *et al.* to estimate apparently robust phylogenetic estimates from an impressively large and equally impressively sparse data set—all collected from existing databases—has important implications for future work on the Tree of Life. This and other studies demonstrate the importance of computational approaches compared with brute-force gene sequencing. In addition, Driskell *et al.*'s approach allows for a more comprehensive sampling of existing data, which enables characterization of evolution across the greatest possible diversity of life. Driskell *et al.* report that more than 100,000 species have at least one molecular sequence archived in public databases. They fail, however, to mention the extreme sampling bias in these databases toward vertebrate animals and green plants. They suggest that the number of these favored species is 6% of those known to science, which is roughly equivalent to the National Science Foundation's estimate of 1.7 million species. However, estimates of the total number of species on Earth (not just those known to science) range from 4 million to 100 million. The computational methodology of Driskell and colleagues may be well suited to future efforts both in terms of taxon sampling and gene sampling for maximizing coverage of the Tree of Life. This represents an extreme departure from the notion of "barcoding" all of life, which emphasizes sequencing one gene for all species (*9*). Both approaches rely heavily on having a well-populated database, but the supertree strategy does not have the constraint that these data must be from the same gene. It may turn out that two different databases are needed for these two distinct purposes (that is, establishing relationships versus diagnosing species).



Building the Tree of Life. A current view of the Tree of Life (*7*). Information is biased toward vertebrate animals and vascular plants (the thick branches); lesser-known groups such as bacteria, fungi, and protists are largely underrepresented. Also shown are species known to science (green leaves), extinct species (leaf litter, brown), endangered species (falling leaves), and species for which "barcode" information is available (red leaves).

If the supertree approach establishes the trunk and thick branches of the Tree of Life, then perhaps the barcoding approach is more appropriate for discerning the twigs and leaves of the tree (see the figure). Currently, most attention has been focused on the trunk at the expense of the leaves. However, the leaves are dropping quickly. We are losing 27,000 species each year while only describing 18,000 new species (*10*). The Driskell *et al.* study provides hope for combining diverse and sparse data sets collected from both leaf and trunk areas of the Tree of Life to provide a robust estimate of this tree, but this should in no way undermine efforts to characterize as many leaves as possible before they hit the ground. Future applications of Driskell *et al.*'s computational method and verification of its performance compared with computer-simulated known phylogenetic histories and empirically "known" histories will provide further insights into the generality of mining our genetic databases to assemble the Tree of Life.

References and Notes

1. M. Pagel, *Nature* **401**, 877 (1999).
2. A new NSF program funds computational approaches for "assembling the Tree of Life" (AToL). Total AToL program funding is $13 million for fiscal year 2004. NSF, *Assembling the Tree of Life: Program Solicitation NSF 04-526* (www.nsf.gov/pubs/2004/ nsf04526/nsf04526.pdf).
3. A. C. Driskell *et al.*, *Science* **306**, 1172 (2004).
4. M. J. Sanderson *et al.*, *Trends Ecol. Evol.* **13**, 105 (1998).
5. J. Wiens, *Syst. Biol.* **52**, 528 (2003).
6. J. W. Thornton, R. DeSalle, *Annu. Rev. Genomics Hum. Genet.* **1**, 41 (2000).
7. M. C. Rivera, J. A. Lake, *Nature* **431**, 152 (2004).
8. W. Doolittle, *Science* **284**, 2124 (1999).
9. P. D. N. Herbert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14812 (2004).
10. E. O. Wilson, *The Diversity of Life* (Norton, New York, 1992).

Biodiversity Conservation and the Eradication of Poverty

William. M. Adams,[1] Ros Aveling,[2] Dan Brockington,[3] Barney Dickson,[2] Jo Elliott,[5] Jon Hutton,[4] Dilys Roe,[6] Bhaskar Vira,[1] William Wolmer[7]

It is widely accepted that biodiversity loss and poverty are linked problems and that conservation and poverty reduction should be tackled together. However, success with integrated strategies is elusive. There is sharp debate about the social impacts of conservation programs and the success of community-based approaches to conservation. Clear conceptual frameworks are needed if policies in these two areas are to be combined. We review the links between poverty alleviation and biodiversity conservation and present a conceptual typology of these relationships.

Biodiversity conservation scientists face a dilemma. There is increasing concern that global efforts to maintain biodiversity are in conflict with those to reduce poverty (*1*). The decline of populations, extinction of species, and habitat transformation demand urgent action (*2*). The leading response to these threats since the late 19th century has been the creation of protected areas (*3*). Technical capacity to design effective protected-area systems is increasing (*4*), allowing the identification of coverage and remaining gaps in the international protected-area system (*5*). This, combined with positive assessments of the effectiveness of protected areas is encouraging the consolidation and expansion of the network of protected areas (*6*). The 2004 World Database on Protected Areas includes over 105,000 sites covering an area of 19.7 million km (*2*, *7*). The establishment and effective management of a global series of protected areas was a key element of the 7th Conference of the Parties (COP) to the Convention on Biological Diversity (CBD) in 2004 (*8*).

The problem with this strategy is that its impacts on poverty are often negative. The creation of protected areas causes the foreclosure of future land use options, with potentially significant economic opportunity costs (*9*). The creation of protected areas can have substantial negative impacts on local people. The eviction of former occupiers or right holders in land or resources can cause the exacerbation of poverty, as well as contravention of legal or human rights (*10*–*14*). Globally, it is recognized that the costs of biodiversity conservation are not distributed in proportion to their benefits (*15*). Typically, many of the costs of protected areas in poor biodiverse countries are paid by local people (*16*). The 7th CBD COP called for an assessment of "the economic and sociocultural costs of protected areas (including the cost of livelihood opportunities foregone), and policies to ensure that they are equitably compensated" (*8*). By the start of the 21st century, a remarkable international agreement on the urgency of global poverty elimination had made the relation between biodiversity conservation and poverty reduction an important element of debate about conservation policy (*1*, *13*).

The meaning of poverty may be intuitively obvious, but its measurement is complex. Common definitions are based on monetary (such as per-capita income) or nonmonetary (such as health or mortality) criteria, although broader approaches have been suggested (*17*, *18*). In 1999, 1.2 billion people worldwide had consumption levels below $1 a day and 2.8 billion lived on less than $2 a day (*17*). Poverty is not a static condition, but it is estimated that between 300 and 420 million people live in a state of chronic poverty (always or usually poor) (*19*). The first of the United Nations Millennium Development Goals (MDGs), agreed on in 2000, was to halve, between 1990 and 2015, both the proportion of people whose income is less than $1 a day and the proportion of people who suffer from hunger (*20*).

National poverty reduction strategies are central to attempts to achieve poverty elimination (*21*, *22*). There is a clear need for these to be integrated with national sustainable development strategies (*1*, *23*). The UN MDGs are premised on such integration, with the area of land protected to maintain biological diversity being an indicator of performance against MDG Goal 7 ("to ensure environmental sustainability"). However, the co-listing of poverty elimination and environmental goals does not mean that integrated solutions are possible or that protected areas can contribute to growth and poverty reduction in poor countries. Indeed, the separation by the MDGs of environmental sustainability issues from development goals alarms some observers (*24*). It has even been suggested that the urgent global push for poverty reduction has subsumed or supplanted conservation goals (*1*).

Combining Conservation and Development Goals

The combination of poverty elimination and biodiversity conservation goals has been approached in various ways. The specific problem of the social impacts of protected areas has been recognized by conservation planners for two decades. The principle that the needs of local people should be systematically integrated into protected-area planning was agreed to at the third World Parks Congress in Bali in 1982 (*25*). In 1992, the president of IUCN–The World Conservation Union argued that "if local people do not support protected areas, then protected areas cannot last" (*26*). IUCN's director general now suggests that protected areas should be seen as "islands of biodiversity in an ocean of sustainable human development," with their benefits extending far beyond their boundaries (*27*), but this is still an aspiration. Delegates from the human rights and minority peoples' movements prominently voiced concern at the persistence of such impacts at the fifth World Parks Congress in September 2003 (*28*). There are coherent calls for better understanding of the social impacts of protected areas (*29*, *30*).

Beyond protected areas, the question of whether it is possible to combine poverty elimination and biodiversity conservation relates to the more general debate, familiar

[1]Department of Geography, University of Cambridge, Cambridge, CB2 3EN, UK. [2]Fauna & Flora International, Great Eastern House, Tenison Road, Cambridge, CB1 2RS, UK. [3]School of Geography and the Environment, University of Oxford, Mansfield Road, Oxford, OX1 3TB, UK. [4]Resource Africa, Post Office Box 198, Cambridge, CB3 0TF, UK. [5]Department for International Development, 1 Palace Street, London SW1E 5HE, UK. [6]International Institute for Environment and Development, 3 Endsleigh Street, London WC1H 0DD, UK. [7]Institute of Development Studies, University of Sussex, Falmer, Sussex, BN1 9RE, UK.

to conservation scientists, about the environmental dimensions of development. In the 20th century, the dominant approach was to push for economic growth first and assume that environmental problems (and indeed improved social welfare) could be sorted out later. Economists argue that as economies grow, they can invest in cleaner technologies and less resource-depleting processes: Arguably, an ''environmental Kuznets curve'' can be observed in industrialized and newly industrialized countries, with improvements in factors such as air pollution (*31*). In the 1950s and 1960s, development planners paid scant attention to environmental impacts, whether focusing on poverty elimination, the creation of high-productivity agriculture, or physical infrastructure such as dams or industrialization and the associated problems of pollution (*32*, *33*).

Critics of this technocratic top-down development focused on its environmental and social failures (*33*, *34*). The need to improve the environmental record of development gave rise to the second approach to the environmental aspects of development, in the concept of sustainable development, which underpinned the 1980 World Conservation Strategy document (*33*). As developed since, notably at the World Conference on Environment and Development in Rio de Janeiro in 1992 and the World Summit on Sustainable Development in Johannesburg in 2002, the concept of sustainable development was extended to make explicit reference to justice, equity, and the elimination of poverty. World leaders agreed that biodiversity and resource conservation must be fully integrated into strategies for economic development and are essential elements of sustainable livelihoods at local scales (*35*). It is widely argued that biodiversity underpins the livelihood strategies of the rural poor (*16*). These political and policy insights have been accompanied by the emergence of new academic subfields that offer integrative transdisciplinary insights into social-ecological systems (*36*).

Sceptics point to the large element of wish fulfillment in arguments about the possibility of win-win solutions in sustainable development (*1*, *33*, *37*, *38*). A strong body of opinion, however, maintains that poverty elimination and conservation can happen together. The term ''pro-poor conservation'' has been used to identify conservation strategies that are designed to deliver both poverty reduction and biodiversity protection (*39*, *40*). But is this confidence in win-win solutions justified? Lasting positive outcomes of conservation-with-development projects are elusive (*41*, *42*). Projects that seek to integrate conservation and development have tended to be overambitious and underachieving (*41–44*). Although it is desirable to satisfy the goals of biodiversity and poverty reduction simultaneously, it may only be possible under specific institutional, ecological, and developmental conditions [such as in long-lasting field projects in small human communities in fragile ecosystems (*1*)]. The links between biodiversity and livelihoods, and between conservation and poverty reduction, are dynamic and locally specific (*34*, *45*). In most cases, hard choices will be necessary between goals, with significant costs to one goal or the other. The acceptability of these costs will vary for different organizations and actors.

Diverse Relations Between Conservation and Poverty Reduction

Clarity over the choices between biodiversity conservation and poverty elimination goals is essential. The desire to package projects as delivering win-win solutions plays down the incompatibilities between goals. Equally, exclusive conservation or development goals can be blind to alliances that favor both (*1*). We therefore offer a conceptual typology of the relationships between poverty reduction and conservation in order to promote a clearer understanding of them. The typology presents four different ways of looking at the connections and disconnections between poverty reduction and conservation, reflecting positions in the current debate. It includes both the moral and pragmatic dimensions of arguments for the conservation of biodiversity and the reduction of poverty. Disentangling these makes for clarity.

1) *Poverty and conservation are separate policy realms.* This position sees poverty elimination and conservation as quite different problems comprising distinct sectors of policy concern. Thus, conservation is a legitimate objective that can be pursued independently of any benefits in poverty reduction (and vice versa). Under this position, conservation strategies would focus on the establishment of protected areas or approaches such as direct payments (*46*). If poverty is recognized as an important cause of conservation failure, the response is the designation of further critical biodiverse habitat and the stronger defense of protected areas, rather than the dissipation of scarce conservation resources to maintain diversity across landscapes or in poverty alleviation activities (*37*, *38*). This position sees conservation benefiting poverty reduction indirectly where it secures ecosystem services that yield economic benefits to society, such as enhanced water yields from forested catchments (*47*, *48*). There may also be local opportunities for win-win strategies that combine biodiversity and poverty reduction [such as protected-area tourism arrangements (*49*)]. However, this position holds that trying to combine conservation with poverty reduction everywhere risks misallocating limited conservation resources and compromising biodiversity preservation (*37*, *38*). The key to the success of conservation is the establishment and effective management of a complete global network of protected areas selected because of scientific criteria and owned or legally established by the state or legitimate private owners. Success is measured in terms of biodiversity criteria, not of measures of social development (*6*).

2) *Poverty is a critical constraint on conservation.* This position makes the empirical, pragmatic argument that poverty limits conservation success to a sufficient degree that biodiversity conservation will fail if it does not successfully address poverty elimination. Such a position might be expected in a scenario where poor people were overharvesting wild species, poaching critical species, or colonizing and cultivating biodiverse land, and if the political or economic costs of stopping them (such as by a conventional strict protected-area strategy) were prohibitive. Poverty reduction would be undertaken in this instance simply as a means to achieve more effective conservation. This position holds that to achieve its goal, conservation must provide (and be seen to provide) effective contributions to poverty reduction, including both net benefits to the poor and the avoidance of significant local costs to any social group. Conservation organizations will invest in addressing the poverty of critical protected-area neighbors and actors with the power to disrupt conservation programs. Examples of policy action include classic park outreach strategies (such as service provision to neighboring villages, employment for local people, and participation in park planning processes) and income-generating projects (such as sharing revenue from wildlife tourism in protected areas, integrated conservation-development projects, or the provision of locally acceptable alternatives to lost resources) (*41*, *43*).

3) *Conservation should not compromise poverty reduction.* This position recognizes that conservation agencies have conservation as their primary goal, but it holds that in pursuing that goal they should, at a minimum, not increase poverty or undermine the livelihoods of the poor. This position was adopted at the Fifth World Parks Congress in 2003, but has its critics (*27*). Examples of strategies resulting from this position might include codes of conduct for conservation organizations, social impact assessment of protected areas (*29*, *30*), and the payment of the full local opportunity costs of conservation in protected areas (*50*). Conservation strategies might also seek to generate positive economic benefits for local communities within constraints of biodiversity conservation targets, for example through nonextractive use [such as ecotourism (*49*)] or

harvesting within sustainable limits [such as safari hunting, medicinal products, or biomass products (*51*, *52*)]. This position differs from the empirical claim in position 2 that poor people, if ignored, will undermine conservation. Rather it reflects independent moral and political obligations on conservation agencies to take account of human poverty. It is a claim that recognizes that conservation action can be sustained despite negative social impacts (*53*). It applies even where it is possible to do conservation effectively without benefiting poor people.

4) *Poverty reduction depends on living resource conservation.* This position rests on the empirical claim that financially poor and socially and politically marginalized people depend on living species in biodiverse ecosystems for livelihoods and ecosystem services, and that their livelihoods can be improved through appropriate conservation activities (*33*). Conservation is therefore a tool for achieving poverty reduction, with the sustainable use of natural resources being a foundation of strategies for achieving poverty reduction and social justice. Biodiversity benefits not immediately necessary to this goal are a secondary gain. This position might lead to the rejection of a protected-area strategy because, except under special circumstances (for example, where shares of ecotourism revenues exceeded all other forms of land use), protected areas were unlikely to achieve poverty reduction goals. Alternative approaches would include the sustainable use of living resources to optimize economic return and/or positive impacts on the rural or urban poor (*54*). Examples of policy include conservation programs outside protected areas; for example, to promote the local management of common-pool resources within the constraints of ecological sustainability such as fisheries, wildlife, grazing, or forestry that are targeted at improving the livelihoods of the poor (*54*–*56*). Conservation in response to this position tends toward the maintenance of yields of harvestable species and ecosystems rather than the preservation of biodiversity. Outcomes may deviate to a greater or lesser degree from biodiversity conservation targets. This principle is reflected in the ''ecosystem approach'' adopted by the CBD in 2000 (*57*).

Conclusion

No position outlined here suggests that either the conservation of biodiversity or the elimination of poverty are improper goals. All positions are consistent with the call for conservation organizations to identify and monitor the social impacts of their work, and to take corporate responsibility for operating in a socially accountable manner (*29*). They are also all consistent with the need for poverty alleviation efforts, and wider projects for the development of humankind, to have regard to their demands, or footprint, on the biosphere (*3*, *58*, *59*).

Different agencies (and different individuals) are likely to wish to adopt different positions. For example, differences in thinking about the balance to be struck between poverty reduction and biodiversity conservation underlie different positions in the ''parks versus sustainable use'' debate (*37*, *38*, *54*, *60*). Those advocating strictly enforced protected areas in poor developing countries to guarantee the maintenance of populations of vulnerable species (such as forest primates) are adopting position 1, treating the problems of extinction and poverty as separate. Those advocating programs to tackle the poverty of people living around such parks in order to persuade them not to trespass or hunt are adopting position 2, seeing poverty as a critical constraint on conservation. Those who would seek to increase the flow of revenues from such parks to a level that would fully compensate all stakeholders for associated opportunity costs of the park are adopting position 3, attempting to ensure that conservation does not increase poverty in any way. Those who propose conservation strategies building on the needs of local communities for sustainable harvests of wild species resources, and not necessarily a formally declared protected area at all, are adopting position 4, seeing conservation strategies based on sustainable use primarily as a means to reduce poverty.

Policy that fails to take account of the diverse relationships between conservation needs and the demands of poverty reduction, and the related consumptive demands of the growing world economy, risks failure (*1*). Organizations committed to the preservation of species and those committed to sustainable rural livelihoods based on natural resource use are likely to engage with issues of poverty and biodiversity in very different ways. Their interactions will be facilitated if they can understand their mutual positions. The recognition of different starting points in the way in which biodiversity conservation and poverty elimination goals are prioritized is essential if there is to be success in identifying common ground and differences between biodiversity and development organizations. Such recognition will facilitate the task of those who believe that the goals must be achieved together.

It is premature to abandon attempts to combine conservation and development. The elimination of poverty and the preservation of biodiversity are two distinct objectives. Each may be driven by different moral agendas, but there is considerable overlap in practice.

At the local scale, the policy need is to reconcile the interests of different stakeholders in the management of the natural resources of biodiverse ecosystems (*45*). The larger challenge is to allow human society to meet its potential and share the fruits of economic growth while sustaining a biosphere that not only sustains full ecological functions but retains its living diversity (*3*, *34*).

References and Notes

1. S. E. Sanderson, K. H. Redford, *Oryx* **37**, 1 (2003).
2. S. Palumbi, *Science* **293**, 1786 (2001).
3. W. M. Adams, *Against Extinction: The Story of Conservation* (Earthscan, London, 2004).
4. C. R. Margules, R. L. Pressey, *Nature* **405**, 243 (2000).
5. A. S. Rodrigues *et al.*, *Nature* **428**, 640 (2004).
6. A. G. Brunner, R. E. Gullison, R. E. Rice, G. A. B. da Fonseca, *Science* **291**, 125 (2001).
7. http://sea.unep-wcmc.org/wdbpa/ (6 July 2004).
8. www.biodiv.org (22 February 2004).
9. M. Norton-Griffiths, C. Southey, *Ecol. Econ.* **12**, 125 (1995).
10. D. Brockington, *Fortress Conservation: The Preservation of the Mkomazi Game Reserve, Tanzania* (Currey, Oxford, 2002).
11. C. L. Fortwangler, in *Contested Nature: Promoting International Biodiversity with Social Justice in the Twenty-First Century*, S. R. Brechin, P. R. Wilshusen, C. L. Fortwangler, P. C. West, Eds. (State Univ. of New York Press, Albany, NY, 2003), pp. 25–40.
12. M. Colchester, *Salvaging Nature: Indigenous Peoples, Protected Areas and Biodiversity Conservation* (World Rainforest Movement, Montevideo, 2002).
13. K. Ghimire, M. Pimbert, *Social Change and Conservation* (Earthscan, London, 1997).
14. C. Geisler, R. de Sousa, *Public Adm. Dev.* **21**, 159 (2001).
15. M. Wells, *Ambio* **21**, 237 (1992).
16. D. Roe, J. Elliott, *Oryx* **38**, 137 (2004).
17. World Bank Poverty Net, www.worldbank.org/poverty.
18. A. Sen, *Development as Freedom* (Oxford Univ. Press, Oxford, 2001).
19. www.chronicpoverty.org/chronic_poverty_report_2004.htm (7 July 2004).
20. www.developmentgoals.org (10 March 2004).
21. www.worldbank.org/poverty/strategies/index.htm (19 June 2004).
22. J. Bojo, R. C. Reddy, *Poverty Reduction Strategies and Environment* (World Bank Environment Department Paper 86, World Bank, Washington, DC, 2002).
23. www.un.org/esa/sustdev/natlinfo/nsds/map2002.htm (19 June 2004).
24. D. Roe, in *The Millennium Development Goals: Hitting the Target or Missing the Point?* (IIED, London, 2003), pp. 55–70.
25. A. Phillips, *George Wright Forum* **20**, 8 (2002).
26. S. Ramphal, in *Parks for Life: Report of the IVth World Congress on National Parks and Protected Areas*, J. McNeely, Ed. (IUCN, Gland, Switzerland, 1993), pp. 56–58.
27. A. Steiner, *New Sci.* **180**, 21 (2003).
28. www.iucn.org/themes/wcpa/wpc2003/english/outputs/recommendations.htm (10 March 2004).
29. D. Brockington, K. Schmidt-Soltau, *Oryx* **38**, 140 (2004).
30. C. Geisler, in *Contested Nature: Promoting International Biodiversity Conservation with Social Justice in the Twenty-First Century*, S. R. Brechin, P. R. Wilshusen, C. L. Fortwangler, P. C. West, Eds. (State Univ. of New York Press, Albany, NY), pp. 217–229.
31. World Bank, *World Development Report 1992: Development and the Environment*, (Oxford Univ. Press for the World Bank, New York, 1992).
32. R. F. Dasmann, J. P. Milton, P. H. Freeman, *Ecological Principles for Economic Development* (Wiley, Chichester, UK, 1973).
33. W. M. Adams, *Green Development: Environment and Sustainability in the Third World* (Routledge, London, 2001).
34. S. E. Sanderson, K. H. Redford, *Oryx* **38**, 146 (2004).
35. Livestock and Wildlife Advisory Group, *Wildlife and Poverty Study* (Department for International Development, London, 2002).
36. F. Berkes, *Conserv. Biol.* **18**, 621 (2004).
37. J. Terborgh, *Requiem for Nature* (Island Press, Washington, DC, 1999).
38. J. F. Oates, *Myth and Reality in the Rain Forest: How*

Conservation Strategies Are Failing in West Africa (Univ. of California Press, Berkeley, CA, 1999).
39. IUCN, *Beyond Rhetoric: Putting Conservation to Work for the Poor* (IUCN, Gland, Switzerland, 2002).
40. D. Roe, J. Hutton, J. Elliott, K. Chitepo, M. Saruchera, *Policy Matters* **12**, 52 (2003).
41. D. Hulme, M. W. Murphree, Eds., *African Wildlife & Livelihoods. The Promise and Performance of Community Conservation* (Currey, Oxford, 2001).
42. J. S. Murombedzi, *J. Int. Dev.* **11**, 287 (1999).
43. C. S. Barrett, P. Arcese, *World Dev.* **23**, 1073 (1995).
44. M. Wells, K. Brandon, *People and Parks: Linking Protected Areas with Local Communities* (World Bank, Washington, DC, 1992).
45. T. Kepe, M. Saruchera, W. Whande, *Oryx* **38**, 143 (2004).
46. P. J. Ferraro, *Science* **298**, 1718 (2002).
47. A. Balmford *et al.*, *Science* **297**, 950 (2002).
48. G. C. Daily, Ed., *Nature's Services: Societal Dependence on Natural Ecosystems* (Island Press, Washington, DC, 1997).
49. S. Gösling, *Ecol. Econ.* **29**, 303 (1999).
50. A. Balmford, T. Whitten, *Oryx* **37**, 238 (2003).
51. B. M. Campbell, M. K. Luckert, *Uncovering the Hidden Harvest: Valuation Methods for Woodland and Forest Resources* (Earthscan, London, 2002).
52. D. S. Wilkie, J. F. Carpenter, *Oryx* **3**, 338 (1999).
53. D. Brockington, *Policy Matters* **12**, 22 (2003).
54. J. Hutton, N. Leader-Williams, *Oryx* **37**, 215 (2003).
55. T. Franks, *Issues Nat. Resour. Manage.* **1**, 2 (2003).
56. I. Koziell, J. Saunders, Eds., *Living Off Biodiversity: Exploring the Livelihoods and Biodiversity Issues in Natural Resources Management* (IIED, London, 2001).
57. www.biodiv.org/decisions/default.aspx?m=cop-05 (6 July 2004).
58. M. L. Imhoff *et al.*, *Nature* **429**, 870 (2004).
59. M. Wackernagel, W. Rees, *Our Ecological Footprint: Reducing Human Impact on Earth* (New Society Publishers, Gabriola Island, British Columbia, 1996).
60. P. R. Wilshusen, S. R. Brechin, C. L. Fortwangler, P. C. West, *Soc. Nat. Resources* **15**, 17 (2002).

Pleistocene Brown Bears in the Mid-Continent of North America

Paul Matheus,[1]* James Burns,[2] Jaco Weinstock,[3] Michael Hofreiter[4]

Current biogeographic models hypothesize that brown bears (*Ursus arctos*) migrated from Asia to North America via east Beringia (unglaciated Alaska and Yukon) ~100 to 50 thousand radiocarbon years ago (ka) but did not reach areas south of Beringia until the opening of a mid-continental ice-free corridor (IFC) ~13 to 12 ka (*1*, *2*). This model has been problematic because migration to the mid-continent was blocked by glacial ice only during the relatively short period ~23 to 12 ka, meaning the mid-continent was largely ice-free before the Last Glacial Maximum (LGM), ~24 to 18 ka (*3*–*5*).

Thirty-one brown bear fossils have been radiocarbon-dated in east Beringia (*2*). Three have infinite ages (>53 ka), the rest are <48 ka, and there is a hiatus in their record from ~35 to 21 ka, suggesting brown bears were absent during that interval (*2*). Ancient mitochondrial DNA studies have shown that three clades of brown bears inhabited east Beringia before 35 ka: clades 2c, 3c, and 4 (*2*) (Fig. 1). Populations that recolonized east Beringia after 21 ka all belonged to subclades that were not present before 35 ka (*2*). Clade 4 never reappeared in east Beringia. After 10 ka, the modern phylogeographic structure of North American brown bears became established: clade 3a in western and central Alaska, clade 3b in extreme eastern Alaska and northwestern Canada, and clade 4 in southern Canada and the contiguous United States (*6*). It has been difficult to explain the descendency of bears in the mid-continent if clade 4 has been absent from North America since 35 ka.

Here, we report a well-preserved cranial fragment from a brown bear (Provincial Museum of Alberta no. P98.5.374) collected in fluvial gravels near Edmonton, Alberta (Fig. 1, fig. S1, and table S1). Two accelerator radiocarbon dates on collagen returned ages of 25,210 ± 560 years (AA48743) and 27,410 ± 200 years (OxA-12902). This find indicates that brown bears reached areas south of Beringia well before the postglacial period and before the coalescence of the Laurentide and Cordilleran glaciers.

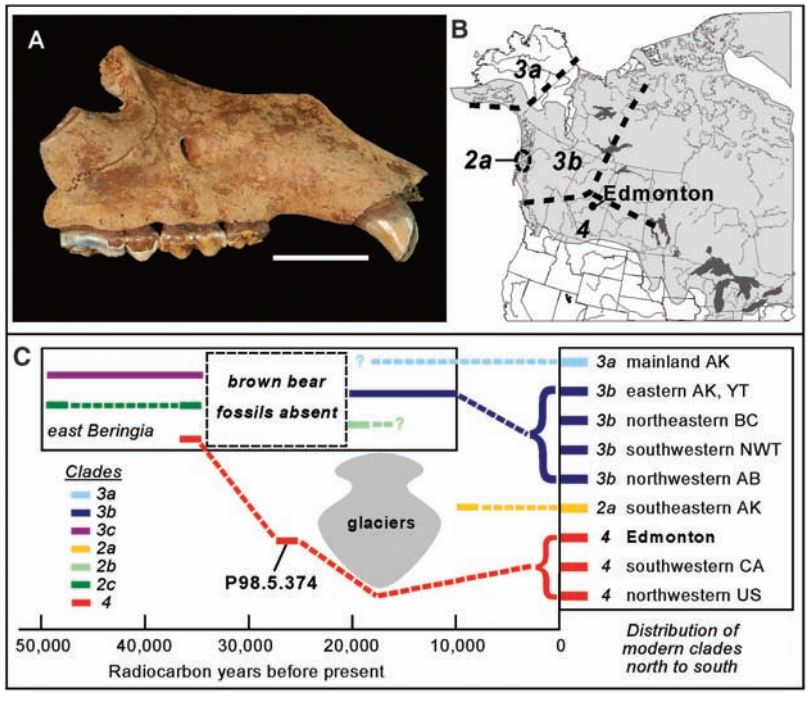


Fig. 1. (**A**) Brown bear cranial fragment P98.5.374 recovered near Edmonton, Alberta. Scale bar, 3 cm. (**B**) The maximum extent of glaciation (gray shading) in northwestern North America during the LGM (*4*) and the present distribution of brown bear clades (*6*). (**C**) Brown bear clades in North America over the past 50,000 years (*2*, *6*). Dashed lines indicate that presence is presumed even though fossils are lacking. The box in the upper left represents populations in Pleistocene east Beringia. AK, Alaska; YT, Yukon Territory; BC, British Columbia; NWT, Northwest Territories; AB, Alberta; CA, Canada; US, United States.

We extracted mitochondrial DNA from ~0.5 g of the root of the specimen's second molar, using established ancient-DNA techniques (*2*, *7*). Two nonoverlapping but highly variable fragments (60 and 135 base pairs) of the control region were amplified, and the sequences obtained show that P98.5.374 belongs to clade 4 (GenBank accession nos. AY796010 and AY796011).

These results challenge the hypothesis that southern brown bears are descended from Beringian populations dispersing through the IFC ~13 to 12 ka, because those dispersers would have belonged to clade 2 or 3. With clade 4 bears inhabiting central Alberta ~26 ka, a more parsimonious model is that clade 4 bears penetrated into southern regions well before the LGM, that they became isolated south of the ice during the LGM, and that modern bears there are descended from in situ populations.

By implication, northern and southern female brown bears in North America have been genetically isolated from each other for at least 35,000 years. Furthermore, the first appearance of brown bears south of the LGM ice margin should not be used to date the earliest availability of a late glacial IFC for human expansion in the New World.

References and Notes

1. B. Kurtén, E. Anderson, *Pleistocene Mammals of North America* (Columbia Univ. Press, New York, 1980).
2. I. Barnes, P. Matheus, B. Shapiro, D. Jensen, A. Cooper, *Science* **295**, 2267 (2002).
3. R. R. Young, J. A. Burns, D. G. Smith, L. D. Arnold, R. B. Rains, *Geology* **22**, 683 (1994).
4. A. S. Dyke *et al.*, *Quat. Sci. Rev.* **21**, 9 (2002).
5. J. A. Burns, *Quat. Int.* **32**, 107 (1996).
6. J. A. Leonard, A. K. Wayne, A. Cooper, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1561 (2000).
7. Materials and methods are available on *Science* Online.
8. Primary DNA sequencing was performed at the Henry Wellcome Ancient Biomolecules Centre, and replication at the Max Planck Institute for Evolutionary Anthropology. Partly supported by the Bureau of Land Management, U.S. Department of the Interior.

Supporting Online Material
www.sciencemag.org/cgi/content/full/306/5699/1150/DC1
Materials and Methods
Fig. S1
Table S1
References and Notes

15 June 2004; accepted 8 September 2004

[1]Alaska Quaternary Center, University of Alaska Fairbanks, Fairbanks, AK, USA. [2]Provincial Museum of Alberta, Edmonton, Canada. [3]Henry Wellcome Ancient Biomolecules Centre, University of Oxford, Oxford, UK. [4]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

*To whom correspondence should be addressed. E-mail: ffpem1@uaf.edu

Odd-Parity Superconductivity in Sr_2RuO_4

K. D. Nelson,[1] Z. Q. Mao,[1]* Y. Maeno,[2,3] Y. Liu[1]†

Phase-sensitive measurements were made on Sr_2RuO_4 to establish unambiguously the odd-parity pairing in this material. The critical current of $Au_{0.5}In_{0.5}$-Sr_2RuO_4 superconducting quantum interference devices prepared on Sr_2RuO_4 single crystals was found to be a maximum for devices with junctions on the same side of the crystal and a minimum for devices with junctions on opposite sides, in the limit of zero magnetic flux; these findings indicate that the phase of the superconducting order parameter in Sr_2RuO_4 changes by π under inversion. This result verifies the odd-parity pairing symmetry and the formation of spin-triplet Cooper pairs in Sr_2RuO_4.

Superconductivity originates from the condensation of Cooper pairs into a macroscopic quantum state described by a single wave function, also known as the superconducting order parameter. Because of the Fermi statistics of the electron, quantum mechanics demands that the parity of the wave function of the Cooper pairs be either even or odd. In the absence of spin-orbit coupling and crystal field, both spin and angular momentum of the Cooper pairs are good quantum numbers. The total spin is 0 in even-parity and 1 in odd-parity superconductors. Consequently, having an even or odd parity in the order parameter corresponds to spin-singlet or spin-triplet superconductivity, respectively. In real crystals with finite spin-orbit coupling and a crystal field, strictly speaking, neither spin nor angular momentum is a good quantum number. In this case, the parity becomes the parameter more appropriate for characterizing the pairing symmetry (*1*). Nonetheless, these two classes of superconductors are still frequently referred to as spin-singlet and spin-triplet superconductors, and furthermore, as *s*- or *d*-wave superconductors for the former and *p*- or *f*-wave superconductors for the latter. Studies of odd-parity superfluid ^{3}He (*2*) and other systems where odd-parity pairing is likely realized, ranging from neutron stars (*2*) to superconductor-ferromagnet microstructures (*3*), have revealed highly unusual physical phenomena, providing strong motivation for the long-standing quest for an example of odd-parity superconductivity.

[1]Department of Physics, Pennsylvania State University, University Park, PA 16802, USA. [2]Department of Physics, Kyoto University, Kyoto 606-8502, Japan. [3]International Innovation Center, Kyoto University, Kyoto 606-8501, Japan.

*Present address: Department of Physics, Tulane University, New Orleans, LA 70118, USA.
†To whom correspondence should be addressed. E-mail: liu@phys.psu.edu

Sr_2RuO_4 (*4*), the only layered-perovskite superconductor containing no Cu (*5*), was predicted to be an odd-parity superconductor (*6*, *7*). A body of experimental evidence for odd-parity superconductivity in Sr_2RuO_4 has been obtained [reviewed in (*8*)]. To date, the strongest evidence for odd-parity superconductivity in Sr_2RuO_4 is found in the nuclear magnetic resonance (NMR) Knight shift result. With a magnetic field applied along the *ab* plane, the Knight shift (proportional to the spin susceptibility after the orbital contribution is subtracted), measured carefully at both O(1) (*9*) and Ru (*10*) sites, was found to be essentially a constant across the superconducting transition temperature T_c of Sr_2RuO_4, a result confirmed by polarized-neutron scattering measurements (*11*) and interpreted (assuming weak spin-orbit coupling in Sr_2RuO_4) as a consequence of spin-triplet pairing in this material.

Although this interpretation of the NMR results is recognized, it is also known that other reasons unrelated to pairing symmetry, such as strong spin-orbit coupling (*12*) or peculiarities of the electronic structure (*13*), can also lead to constant spin susceptibility. For example, vanadium, an *s*-wave superconductor, was found long ago to have a constant Knight shift across T_c (*14*). $YBa_2Cu_3O_{7-\delta}$, a *d*-wave superconductor, was also found to possess a constant Knight shift across T_c when measured at the Cu(2) site with the field applied along the *c* axis (*15*). Therefore, further work was needed to establish unambiguously the odd-parity superconductivity in Sr_2RuO_4 (*8*).

Work on high-T_c superconductors (*16*, *17*) has shown that phase-sensitive measurements, which played a key role in settling the debate on pairing symmetry in the high-T_c arena, will enable the most definitive determination of the parity of the order parameter. In such an experiment, the orientation dependence of the phase of the superconducting order parameter is probed directly through the Josephson effect. In the presence of the spin-orbit coupling, a Josephson current j_s between an odd- and an even-parity superconductor is expected to be

$$j_s \propto \langle \mathrm{Re}(c_{21}s_{21}^*)\ \mathrm{Im}[\Psi^* \mathbf{d}\cdot(\mathbf{n}\times\mathbf{k})]\rangle_{FS} \quad (1)$$

(*18*, *19*), where c_{21} and s_{21} are the spin-orbit and spin-independent parts of the transmission matrix, $\mathbf{k}$ is the wave vector, $\mathbf{n}$ is the unit vector normal to the junction plane, and ψ and $\mathbf{d}$ are functions defining the superconductor order parameters with $\Delta_{even} = i\psi(\mathbf{k})\sigma_y$ for even parity and $\Delta_{odd} = i[\boldsymbol{\sigma}\cdot\mathbf{d}(\mathbf{k})]\sigma_y$ for odd parity (*1*). Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices, * represents complex conjugation, Re and Im are the real and imaginary parts of a complex number, and $\langle\rangle_{FS}$ denotes an appropriate average over the Fermi surface. Under the parity (inverting) operation $\mathbf{k} \rightarrow -\mathbf{k}$, $\psi(\mathbf{k})$ is even but $\mathbf{d}(\mathbf{k})$ is odd.

Interesting consequences are expected from Eq. 1. For example, if Sr_2RuO_4 is an odd-parity superconductor with $\mathbf{d}$ along the *c* axis, $\mathbf{d}\cdot(\mathbf{n}\times\mathbf{k})$, and therefore j_s, will be finite for an in-plane junction ($\mathbf{n}$ and $\mathbf{k}$ in the *ab* plane) and zero for a *c*-axis junction (with $\mathbf{n}\|\mathbf{d}$), a selection rule found previously (*20*). In addition, j_s through two parallel junctions prepared on the same odd-parity superconductor but with normal directions $\mathbf{n}$ and $-\mathbf{n}$, respectively (Fig. 1A), will be out of phase by π (so long as the sample configuration is such that $\mathbf{n}$ is not parallel to $\mathbf{d}$, in which case $j_s = 0$). If such a pair of junctions is incorporated into a superconducting quantum interference device (SQUID), as suggested by Geshkenbein, Larkin, and Barone (GLB) (*21*), the critical current (I_c) of the SQUID would be expected to be modulated by Φ, the flux threading the SQUID, in the form

$$I_c \propto \cos[(\Phi/\Phi_0 + \tfrac{1}{2})\pi] \quad (2)$$

where $\Phi_0 = 2.07 \times 10^{-15}$ Wb is the flux quantum. Therefore, $I_c(\Phi = 0)$ corresponds to a minimum. In contrast, for a conventional SQUID—including a SQUID prepared with both junctions on the same surface of an odd-parity superconductor—$I_c(\Phi = 0)$ corresponds to a maximum. The resistance of the SQUID, measured by applying a current larger than I_c, will exhibit a maximum instead of a minimum at $\Phi = 0$. As noted previously (*22*), however, the GLB experiment does not work for certain sample configurations (e.g., those in which $\mathbf{n}$ is parallel to $\mathbf{d}$).

We pursued the GLB experiment in structures (Fig. 1A) prepared on high-quality single crystals of Sr_2RuO_4 (T_c = 1.4 K)

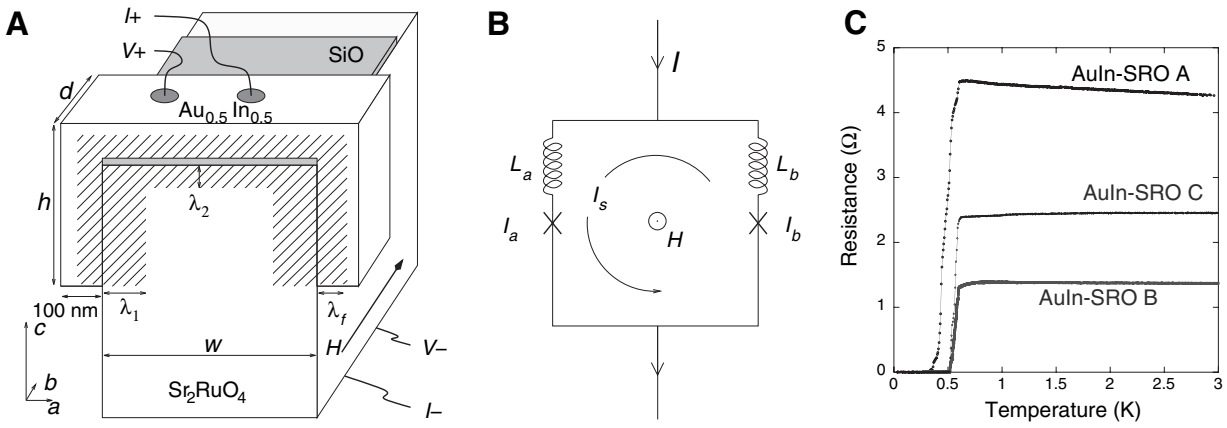


Fig. 1. (**A**) Schematic of $Au_{0.5}In_{0.5}$-Sr_2RuO_4 GLB SQUID with measurement leads. Both junctions in the SQUID are in-plane tunnel junctions to ensure $j_s \neq 0$. The side insulated by the SiO layer may cross the *ab* plane as depicted (samples A and B) or cross a third face perpendicular to the *ab* plane (sample C). The shaded area indicates the flux penetration at $T = 0$, $\lambda_1 \approx 3.7$ μm, and $\lambda_2 \approx 0.18$ μm. For sample C, the flux penetration is to a depth of λ_2 on all three sides. The value of λ_f is not known but may be slightly larger than that of pure In, 0.07 μm. For sample A, $w = 1.05$ mm, $h = 0.5$ mm, and $d = 0.33$ mm; for sample B, $w = 1.15$ mm, $h = 0.4$ mm (left) and 0.6 mm (right), and $d = 0.15$ mm; for sample C, $w = 0.68$ mm, $h = 0.3$ mm, and $d = 0.4$ mm. The SiO layer thickness is 150 nm. (**B**) Equivalent circuit of the SQUID. I_a and I_b are the current on the left and the right side of the SQUID loop, L_a and L_b are effective inductances, I is total current, and I_s is circulating current. (**C**) Sample resistance R as a function of T for three GLB SQUIDs. A smooth $R(T)$ across the T_c of Sr_2RuO_4 (around 1.4 K) indicates that $R(T)$ is dominated by the tunnel barrier.

prepared by the floating zone method. This required us to overcome several technical difficulties (*23*). First, a suitable conventional superconductor had to be identified. Obtaining a SQUID with minimal asymmetry and self-inductance (see below) requires thermal evaporation of the conventional superconductor onto mechanically polished single-crystal surfaces. However, few elemental or alloy superconductors were found to form a Josephson junction with Sr_2RuO_4. Alloy $Au_{1-x}In_x$ prepared by interdiffusion was the only material we found with which Josephson coupling to bulk Sr_2RuO_4 could be achieved with the above requirements. Also, the presence of a minority phase in Sr_2RuO_4—the Ru inclusions referred to as the 3 K phase of Sr_2RuO_4 (*24*)—had to be dealt with carefully. These inclusions are known to form in all single-crystal rods during the crystal growth. Whereas the T_c for the bulk Ru is 0.5 K, the T_c of the Ru inclusions embedded in the bulk Sr_2RuO_4 crystal can be as high as 3 K. The presence of these Ru inclusions at the junction interface could potentially lead to unwanted complications (*23*). We found that close to the surface of a single-crystal rod, there existed areas free of Ru inclusions. Therefore, SQUID samples were prepared on these areas. The presence of Ru inclusions at the junction interface was checked by examining resistance as a function of temperature to see whether any features were present between 1.5 and 3 K (*23*).

The current-voltage characteristics of a Ru-free GLB SQUID are shown in Fig. 2. Josephson coupling marked by a finite, zero-voltage supercurrent is apparent. However, the current-voltage characteristics are asymmetric, which suggests that the two junctions in the SQUID are not identical. The critical temperature of the SQUID determined from $I_c(T = T_c) = 0$ is consistent with that of $Au_{0.5}In_{0.5}$ (*23*). Furthermore, quantum interference patterns in both the critical current and sample resistance were observed (Fig. 3). The period in magnetic field (H) was found to be 8.5 ± 0.4 mG for sample A, 15.7 ± 0.4 mG for sample B, and 98 ± 2 mG for sample C. The direction of the c axis for samples A and B is as indicated in Fig. 1A, whereas for sample C it is pointing out of the page. On the basis of the relevant lengths associated with these samples, and taking into account that the areas of the two junctions in the SQUID are large and that the phase change across the two junctions is position-dependent, the effective area of the SQUID loop (A_{eff}) is 2.1×10^{-5} cm^2 for sample A and 2.2×10^{-5} cm^2 for sample B, which is expected to yield $\Delta H = \Phi_0/A_{eff} = 9$ mG for both junctions. For sample C, A_{eff} is 1.8×10^{-6} cm^2 with the expected $\Delta H = 117$ mG. We attribute the difference between the expected and measured values of ΔH to the possible local variations of the penetration depths and the misalignment of the samples with respect to the applied field.

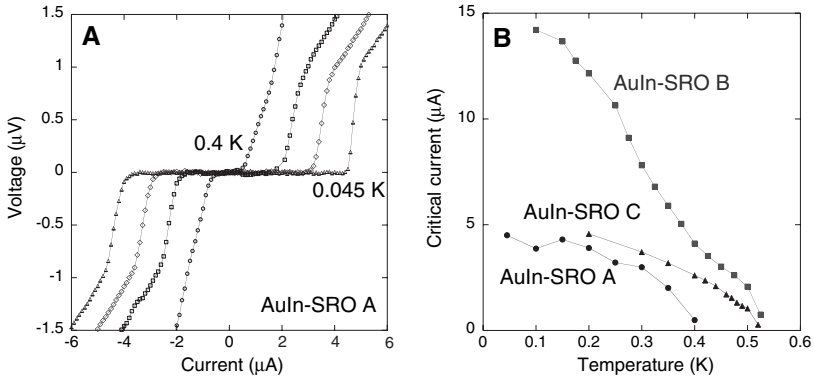


Fig. 2. (**A**) Current-voltage curves of sample A at several temperatures. $T = 0.045$ K (triangles), 0.250 K (diamonds), 0.350 K (squares), and 0.400 K (circles), showing a zero-voltage supercurrent. The asymmetry with respect to the current reversal is due to the presence of a circulating current. (**B**) The temperature dependence of the critical current, $I_c(T)$, determined from a particular (positive) current direction for all three samples.

Two important issues must be addressed before we can show whether $I_c(\Phi = 0)$ corresponds to a minimum or maximum, and therefore whether Sr_2RuO_4 is an odd-parity superconductor. The total flux Φ is given by $\Phi = \Phi_{app} + \Phi_{ind} + \Phi_{trap} + \Phi_{bkgd}$, where Φ_{ind} is the induced flux, Φ_{app} ($= A_{eff}H$) is the applied flux, Φ_{trap} is the trapped flux, and Φ_{bkgd} is the flux due to any background field. To find out what values of Φ_{app} would result in $\Phi = 0$, it is necessary

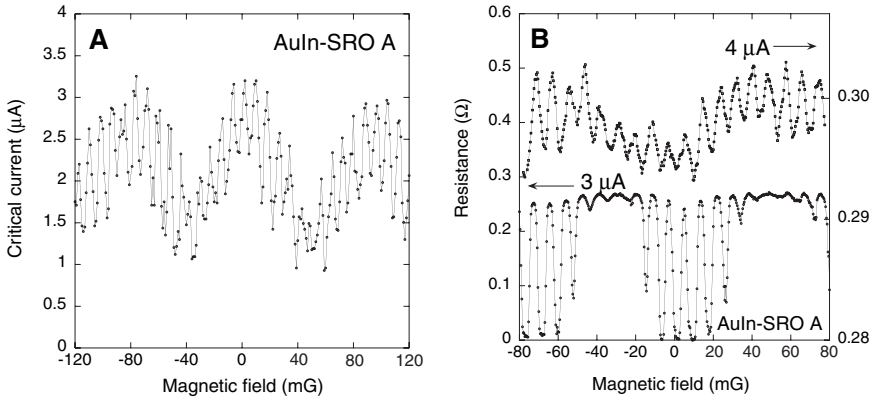


Fig. 3. Examples of quantum oscillations in $I_c(H)$ (**A**) and $R(H)$ (**B**). Values of the current used to measure $R(H)$ are indicated. Both sets of data were taken on sample A at $T = 0.30$ K.

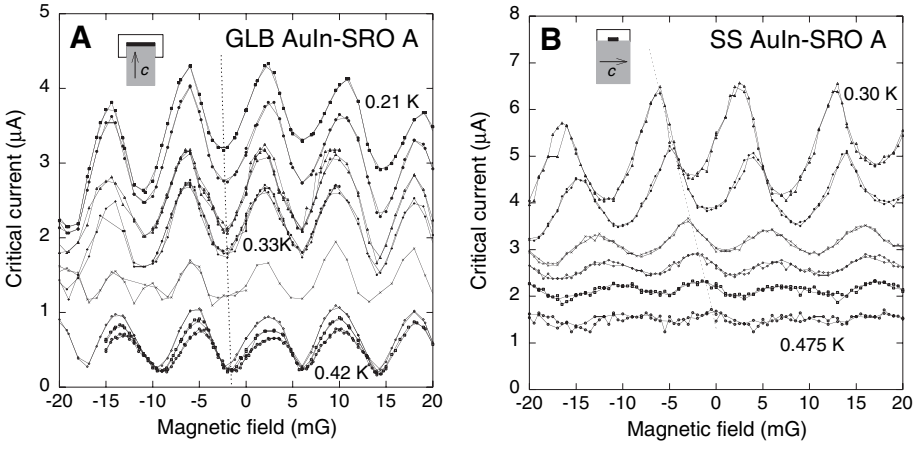


Fig. 4. (**A**) $I_c(H)$ for GLB sample A (opposite-side GLB SQUID) taken close to T_c of the SQUID. From top to bottom, $T = 0.210$, 0.250, 0.300, 0.330, 0.360, 0.380, 0.400, and 0.420 K. A dashed line is used to highlight the shift of the minimum in $I_c(\Phi_{app})$ toward $\Phi_{app} = 0$ as T approaches T_c. (**B**) $I_c(H)$ for same-side (SS) sample A, showing a maximum $I_c(\Phi_{app})$ as T approaches T_c. From top to bottom, $T = 0.300$, 0.350, 0.400, 0.425, 0.450, and 0.475 K. For each curve, the field was swept in both directions to ensure that flux was not trapped during the sweep.

to determine the value of Φ_{ind} and to avoid or minimize Φ_{trap} and Φ_{bkgd} (*25*). The induced flux $\Phi_{ind} = LI_s = L_aI_a - L_bI_b$ (Fig. 1B) is intrinsic for SQUIDs with large areas and an unavoidable asymmetry between the two junctions in the SQUID, whereas the trapped flux Φ_{trap} is typically a by-product of the cooling and measurements (*26*). As shown in the high-T_c work (*16*), the presence of Φ_{trap} in an ordinary SQUID could mimic the behavior of an unconventional SQUID.

The following steps were taken to address these issues: We sought to reduce L and therefore Φ_{ind} by minimizing the width of the SQUID w, which was unfortunately limited by the tendency of the crystal to cleave during the mechanical polishing. Using thin films of Sr_2RuO_4 could, in principle, help; however, even the highest quality films of Sr_2RuO_4 grown so far are not yet superconducting (*27*). The two junctions in the SQUID were prepared simultaneously by thermally evaporating $Au_{0.5}In_{0.5}$ (*23*), rather than pressing the bulk In wire as done previously (*20*) onto the polished Sr_2RuO_4 surface, to obtain a sample with minimal inductance and asymmetry. Furthermore, we measured systematically the $I_c(\Phi_{app}, T)$ of selected SQUID samples with minimal asymmetry ($I_a \approx I_b$ and $L_a \approx L_b$) as T was raised to T_c. In this case, I_a and I_b would each decrease to zero, so that $LI_s = L_aI_a - L_bI_b \rightarrow 0$ and $\Phi = \Phi_{app} + \Phi_{ind} \rightarrow \Phi_{app}$, if $\Phi_{trap} = 0$.

To avoid Φ_{trap}, we carefully shielded the SQUID magnetically (*23*) to minimize the background field (<1 mG) and cooled it down slowly to initialize a Φ_{trap}-free state, as done previously (*28*). The absence of Φ_{trap} was ensured by verifying that the envelope of $I_c(\Phi)$ was symmetric (*28*). We followed a procedure in which the sample, the magnet coil, and the magnetic shielding enclosure were warmed to 14 K or higher and cooled back down slowly (with a uniform rate of $dT/dt = 2$ K/hour) in zero field repeatedly to obtain symmetric $I_c(H)$ traces. In addition, we found that applying a large field appeared to lead to an asymmetric $I_c(\Phi)$, a sign that Φ_{trap} was present. The applied field was therefore kept below a relatively small value. Finally, Φ_{bkgd} is minimized by the good magnetic shield.

These strategies to minimize Φ_{ind}, Φ_{trap}, and Φ_{bkgd} appeared to work fairly well: Nearly symmetric patterns of $I_c(H)$, taken at different temperatures, were obtained on three fairly symmetric GLB SQUID samples. The actual residual field contributing to the residual flux ($\Phi_{ind} + \Phi_{trap}$) could be inferred from the position of the center of the envelope in $I_c(H)$ (1 to 2 mG for sample A; see Fig. 4A). For such nearly symmetric patterns, the overall $I_c(H, T)$ decreased as the temperature increased to T_c, accompanied by shifts of the patterns that were clearly due to the decrease in Φ_{ind} (Fig. 4A). As a self-consistency check, we estimated the values of L from the shift in $I_c(H, T)$ and from the sample dimensions and found a reasonable agreement (*23*).

The detailed evolution of the $I_c(H, T)$ shown in Fig. 4A clearly shows that $I_c(\Phi)$ is a minimum at $\Phi = 0$. Other GLB SQUIDs, samples B and C, showed essentially the same behavior (*23*). On the other hand, control samples with both junctions of the SQUID on the same side (SS) of the crystal (perpendicular to the *ab* plane) were found to possess an $I_c(\Phi)$ interference pattern with a maximum at $\Phi = 0$ (Fig. 4B), as expected given that the phase of the order parameter is the same at both junctions. We also prepared another type of control sample in a corner-junction configuration, which was used in the high-T_c work (*16*). In the interference pattern for such a corner junction, the critical current at $\Phi = 0$ is a maximum for an *s*-wave, a minimum for a *d*-wave, and neither a minimum nor a maximum for a *p*-wave superconductor. Experimentally, a $Au_{0.5}In_{0.5}$-Sr_2RuO_4 corner junction was found to show an interference pattern consistent with *p*-wave pairing (*23*).

Our observations indicate that the phase of the superconducting order parameter in Sr_2RuO_4 changes by π after inversion. This establishment of the odd-parity superconductivity in Sr_2RuO_4 by phase-sensitive experiments constrains the interpretation of other pairing symmetry–related experiments to the odd-parity scenario, which in turn provides a solid starting point to determine the detailed **k**-dependence of the order parameter. This will help in resolving, for example, the "*p*-versus *f*-wave" issue (*8*). The tetragonal crystal symmetry of Sr_2RuO_4 allows only several odd-parity states (*29*). The result of the selection rule in Josephson coupling (*20*), which can now be interpreted with considerable certainty to indicate that **d** is along the *c* axis, reduces the number of allowed odd-parity states even further. The establishment

of an odd-parity superconductor may also provide an arena to explore its possible applications. For example, all electronic spins in superconducting, nonmagnetic Sr_2RuO_4 are condensed into the *ab* plane [because **d** is along the *c* axis (*29*)], a fact possibly useful for novel devices.

References and Notes

1. G. E. Volovik, L. P. Gor'kov, *Sov. Phys. JETP* **61**, 843 (1985).
2. D. R. Tilley, J. Tilley, *Superfluidity and Superconductivity* (A. Hilger, Bristol, UK, ed. 2, 1986).
3. F. S. Bergeret, K. B. Eftov, A. I. Larkin, *Phys. Rev. B* **62**, 11872 (2000).
4. S. N. Ruddlesden, P. Popper, *Acta Crystallogr.* **10**, 538 (1957).
5. Y. Maeno *et al.*, *Nature* **372**, 532 (1994).
6. T. M. Rice, M. Sigrist, *J. Phys. Condens. Matter* **7**, L643 (1995).
7. G. Baskaran, *Physica B* **223–224**, 490 (1996).
8. A. P. Mackenzie, Y. Maeno, *Rev. Mod. Phys.* **75**, 657 (2003).
9. K. Ishida *et al.*, *Nature* **396**, 658 (1998).
10. K. Ishida *et al.*, *Phys. Rev. B* **63**, 060507 (2001).
11. J. A. Duffy *et al.*, *Phys. Rev. Lett.* **85**, 5412 (2000).
12. P. W. Anderson, *Phys. Rev. Lett.* **3**, 325 (1959).
13. D. E. MacLaughlin, *Solid State Phys.* **31**, 2 (1976).
14. R. J. Noer, W. D. Knight, *Rev. Mod. Phys.* **36**, 177 (1964).
15. S. E. Barrett *et al.*, *Phys. Rev. B* **41**, 6283 (1990).
16. D. J. van Harlingen, *Rev. Mod. Phys.* **67**, 515 (1995).
17. C. C. Tsuei, J. R. Kirtley, *Rev. Mod. Phys.* **72**, 969 (2000).
18. V. B. Geshkenbein, A. I. Larkin, *JETP Lett.* **43**, 395 (1986).
19. A. Millis, D. Rainer, J. A. Sauls, *Phys. Rev. B* **38**, 4504 (1988).
20. R. Jin, Y. Liu, Z. Mao, Y. Maeno, *Europhys. Lett.* **51**, 341 (2000).
21. V. B. Geshkenbein, A. I. Larkin, A. Barone, *Phys. Rev. B* **36**, 235 (1987).
22. A. J. Leggett, *Philos. Mag. B* **74**, 509 (1996).
23. See supporting data on *Science* Online.
24. Y. Maeno *et al.*, *Phys. Rev. Lett.* **81**, 3765 (1998).
25. This issue was encountered in the pioneering work of SQUID-based phase-sensitive experiments on high-T_c superconductors (*30*). The SQUID resistance $R(H, I)$, measured at $I > I_c$ (in which case finite Φ_{ind} was necessarily included), was extrapolated to $I = 0$ and the minimum position for $R(H_{min}, 0)$ was inferred. The extrapolation was avoided in the elegant (single) corner-junction experiment.
26. The requirement for the superconducting phase to be single-valued in the sample leads to the constraint $2\pi m = \Phi_1 - \Phi_2 + (2\pi/\Phi_0)(\Phi_{ext} + \Phi_{ind} + \Phi_{bkgd} + \Phi_{trap})$, where m is an integer (or more likely, 0) and Φ_1 and Φ_2 are the phase differences at the two junctions in the SQUID, respectively. Φ_1 and Φ_2 can adjust themselves to accommodate any arbitrary, not necessarily quantized, amount of Φ_{trap}.
27. M. A. Zurbuchen *et al.*, *Appl. Phys. Lett.* **78**, 2351 (2001).
28. B. Chesca *et al.*, *Appl. Phys. Lett.* **76**, 912 (2000).
29. V. P. Mineev, K. V. Samokhin, *Introduction to Unconventional Superconductivity* (Gordon and Beach, Amsterdam, 1999).
30. A. Wollman, D. J. van Harlingen, W. C. Lee, D. M. Ginsberg, A. J. Leggett, *Phys. Rev. Lett.* **71**, 2134 (1993).

31. Supported by NSF grants DMR-9974327 and DMR 0202534, Japan Society for the Promotion of Science and Ministry of Education, Culture, Sports, Science, and Technology grants-in-aid, and the 21st Century Center of Excellence program. We thank V. B. Geshkenbein, J. Kirtley, C.-C. Tsuei, D. F. Agterberg, J. Banavar, J. K. Jain, A. J. Leggett, J. A. Sauls, M. Sigrist, D. J. van Harlingen, and S.-K. Yip for useful discussions, and C. Andreou, R. Jin, H. Wang, D. Schlom, and H. Yaguchi for discussions and assistance.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1151/DC1
Materials and Methods
Figs. S1 to S9
References

10 August 2004; accepted 21 September 2004

Disorder-Sensitive Phase Formation Linked to Metamagnetic Quantum Criticality

S. A. Grigera,[1]* P. Gegenwart,[1,2] R. A. Borzi,[1] F. Weickert,[2] A. J. Schofield,[3] R. S. Perry,[1,4,5] T. Tayama,[6] T. Sakakibara,[6] Y. Maeno,[4,5] A. G. Green,[1] A. P. Mackenzie[1]*

Condensed systems of strongly interacting electrons are ideal for the study of quantum complexity. It has become possible to promote the formation of new quantum phases by explicitly tuning systems toward special low-temperature quantum critical points. So far, the clearest examples have been appearances of superconductivity near pressure-tuned antiferromagnetic quantum critical points. We present experimental evidence for the formation of a non-superconducting phase in the vicinity of a magnetic field–tuned quantum critical point in ultrapure crystals of the ruthenate metal $Sr_3Ru_2O_7$, and we discuss the possibility that the observed phase is due to a spin-dependent symmetry-breaking Fermi surface distortion.

The field of quantum criticality continues to attract widespread theoretical and experimental attention because of its importance to the global effort to understand the behavior of correlated electron systems (*1–4*). It gives the rare opportunity for controlled study of the collective behavior that arises in many-body quantum physics in the presence of strong interactions. Much of the interest has focused on the effects that quantum critical fluctuations have on itinerant systems, notably the link between their strength and the breakdown of Landau's Fermi liquid theory (*5*). Quantum criticality is now understood not to be a trivial extension of the classical case; quantum critical fluctuations can have surprising strength and subtlety, including important mode-mode interaction terms (*3*, *4*).

Recently several examples have been discovered of a phenomenon that is potentially even more exciting, namely the use of quantum critical points to create regions of phase space in which systems are highly susceptible to the emergence of new ordered phases. Qualitatively, the high susceptibility to novel phase formation is thought to arise because tuning a thermal phase transition toward zero temperature flattens the free energy landscape, causing competing low-temperature phases to become nearly degenerate. The characteristics of the quantum critical fluctuations associated with particular quantum critical points (QCPs) can then tip the balance in favor of a particular new phase. A notable example is the formation of unconventional superconductivity in the vicinity of antiferromagnetic QCPs (*6*). In this case, the diverging spin fluctuations are thought to provide the effective bosons for the Cooper pair binding.

Superconductivity is not the only form of low-temperature order that can be adopted by correlated electron systems; the fractional quantum Hall state is a spectacular example of the subtle quantum self-organization that is possible (*7*). A particularly interesting issue is whether forms of magnetically tuned QCPs can be used to explicitly promote the formation of other new types of quantum order (*8*) in the special situation in which the magnetic field disfavors superconducting ground states. We demonstrate the formation of a previously unknown, non-superconducting phase in an applied magnetic field of 8 T, near a metamagnetic QCP in $Sr_3Ru_2O_7$. The experiments place strong constraints on the order parameter of this phase, which appear to be satisfied by invoking a spin-dependent Fermi-surface instability, the possibility of which was first discussed theoretically nearly half a century ago by Pomeranchuk in phenomenological calculations using Landau's Fermi liquid theory (*9*).

[1]School of Physics and Astronomy, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9SS, Scotland. [2]Max-Planck Institute for Chemical Physics of Solids, D-01187 Dresden, Germany. [3]School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [4]International Innovation Center, Kyoto University, Kyoto 606-8501, Japan. [5]Department of Physics, Kyoto University, Kyoto 606-8502, Japan. [6]Institute of Solid-State Physics, University of Tokyo, Kashiwa, Chiba 2778581, Japan.

*To whom correspondence should be addressed. E-mail: sag2@st-and.ac.uk; apm9@st-and.ac.uk

$Sr_3Ru_2O_7$ is the bilayer member of the Ruddlesden-Popper series of layered perovskite ruthenates whose single-layer member, Sr_2RuO_4, is a well-studied unconventional superconductor (*10*). The growth of high-quality single crystals in an image furnace revealed it to be a strongly enhanced paramagnet in zero field, displaying an itinerant metamagnetic transition in applied fields in the range from 4.9 T (field parallel to *ab* plane) to 7.9 T [field parallel to *c* (*11*, *12*)]. Good evidence has been acquired that, for fields parallel to *ab*, the metamagnetism is due to a line of first-order phase transitions terminating in a finite temperature critical point. As the field is rotated away from the *ab* planes, this critical end point is tuned downward in temperature, becoming quantum critical at an angle less than 10° from *c* (*12*–*14*). The crystals on which the work of (*12*–*14*) was performed were of sufficiently high purity (residual resistivity $\rho_{res} \sim 3$ μΩ cm) that disorder might not have been expected to play an important role in determining their properties. However, a further improvement of nearly an order of magnitude to $\rho_{res} \sim 0.4$ μΩ cm (*15*) revealed substantial changes in the observed behavior. Away from the QCP, the properties are indeed essentially independent of purity, but in its immediate vicinity a large and striking peak was observed in the resistivity, in a region bounded at 100 mK by two first-order magnetic phase transitions (*16*). Those observations opened the possibility of the formation of a non-superconducting phase in the vicinity of the metamagnetic QCP in $Sr_3Ru_2O_7$.

We show the magnetic field (H) dependence (for H // c) of the resistivity of a high-purity single crystal of $Sr_3Ru_2O_7$ as temperature increased from 0.1 to 1.3 K (Fig. 1). At 0.1 K, the large, steep-sided feature is similar to that reported and discussed in (*16*). By 1.3 K, the observed resistivity behavior is unsurprising, with a pronounced thermal broadening. The surprising results come at intermediate temperatures. At all temperatures between 0.1 and 1.1 K, the anticipated broadening is seen outside the field range from 7.8 to 8.1 T; but once that field range is entered, ρ appears to "lock in" to the field and temperature dependence seen at the lowest temperature. The value of ρ at which this occurs is temperature-dependent, but there is almost no temperature dependence to the field at which it occurs or to its field dependence once it has locked in to the anomalous resistivity.

The data of Fig. 1 suggest the presence of a low-temperature phase in which the strong temperature-dependent (inelastic) scattering associated with quantum critical fluctuations is cut off at well-defined boundaries and replaced by a large and essentially temperature-independent (elastic) scattering mechanism. It is not, however, compelling thermodynamic evidence of the kind necessary to support the postulate of a phase change. Such thermodynamic signatures are provided by the susceptibility, magnetostriction, thermal expansion, and magnetization. A sample comparison of alternating current (ac) susceptibility (χ) and linear magnetostriction (λ) is shown (Fig. 2). There is a clear correlation between the real part of χ and λ, indicating a strong magnetostructural coupling. A low-field peak at about 7.5 T, which is probably a crossover, is followed by two sharp peaks at fields corresponding precisely to the steep walls in the resistivity of Fig. 1. As discussed in (*13*, *16*), the appearance of a dissipative peak in χ'' indicates that each of these two peaks corresponds to a first-order phase boundary at 100 mK (*17*). All peaks in both χ' and λ are positive, showing that both sample moment and length increase at each transition.

Temperature scans of thermodynamic quantities also show well-defined features at the boundaries of the anomalous region seen in the resistivity, for example, the temperature dependence of the magnetization in a fixed field of 7.9 T (Fig. 2, inset). At temperatures above 1.1 K, the data show upward curvature consistent with divergent fluctuations on the approach to a QCP, but this is then cut off at a well-defined kink, followed by a much weaker temperature dependence. The kink is seen only on entry to the anomalous region; scans at fields outside this range remain smooth to the lowest temperature reached (50 mK).

The data shown in Fig. 2 represent only a small subset of the experiments performed. A collation of all the relevant information is presented (Fig. 3), which also includes points

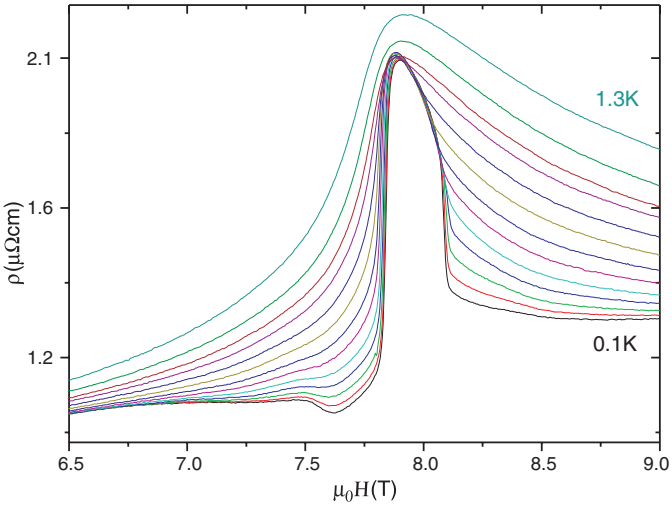


Fig. 1. The resistivity ρ of very pure single–crystal $Sr_3Ru_2O_7$ as a function of magnetic field (applied parallel to c) at a series of temperatures between 0.1 and 1.3 K in steps of 100 mK. Increasing the temperature broadens and increases the resistivity outside the field range of the pronounced central feature, but within that field range and below 1.1 K, it locks in to the peak. The value of ρ at which this occurs is strongly temperature-dependent, but both the lock-in field and the variation of ρ through the peak are strikingly independent of temperature.

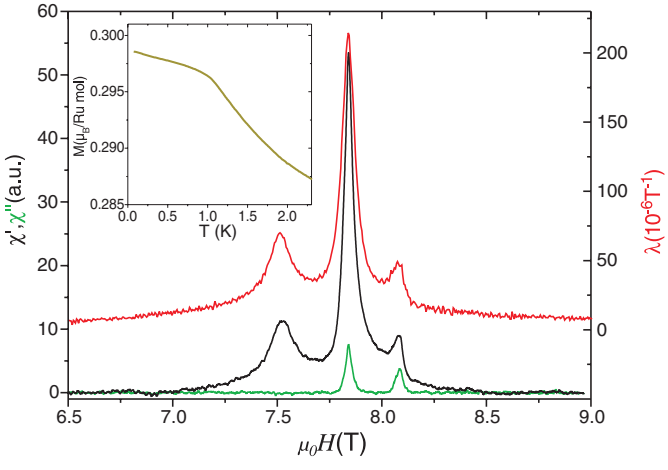


Fig. 2. Sample ac magnetic susceptibility (χ) and linear magnetostriction (λ) data at 100 mK. The susceptibility was measured at a low frequency of 17 Hz to avoid excessive finite-frequency effects (*13*). Both the main field and the small ac field of 3×10^{-5} T root mean square were applied parallel to *c*, and $\lambda = d(\Delta c/c)/dB$. The close similarity between λ and χ' demonstrates a strong magnetostructural coupling, whereas the peaks in χ'' accompanying the two higher-field peaks in λ and χ' identify two first-order phase transitions [also (*17*)]. a.u., arbitrary units. **(Inset)** Sample dc magnetization (M) data. As the temperature is lowered at a fixed field (in this case 7.9 T), the incipient divergence of fluctuations seen in M is cut off at a clearly identifiable kink at the same temperature as that of the top of the peak seen in Fig. 1.

obtained from plotting the positions of peaks in the thermal expansion α. Data obtained from five different crystals of similar purity studied independently in four separate laboratories yield consistent information. Surrounding the temperature and field at which a QCP had been identified in more disordered samples is a region in the (H,T) plane enclosed by well-defined phase boundaries. At sufficiently low temperatures these are first-order, but they then change to second-order as indicated (Fig. 3, red arrows) (*18*).

The existence of this new phase is the central experimental result of our paper. The anomalous transport properties that accompany it are strongly purity-dependent (*14*, *16*); it can be defined with the clarity shown (Fig. 3) only in the very best samples with $\rho_{res} < 1$ μΩ cm. A strong purity dependence like this is characteristic of some of the best-known ordered phases in itinerant interacting electron systems such as the fractional quantum Hall state (*7*) and unconventional superconductivity (*10*). Several key pieces of experimental evidence further indicate that the previously unknown phase is intrinsically linked with the existence of the previously identified metamagnetic QCP. First, as stated above, it encloses the characteristic field of that QCP as estimated either by direct measurement on more disordered samples or by extrapolation of transport and magnetic measurements taken at temperature $T > 1$ K on the highest purity crystals. Second, we draw attention to the relative orientation of the first-order phase boundaries (Fig. 3). These are seen to have opposite slopes and curvatures with respect to field variation, similar to the quantum critical contours that can be calculated by using a Hertz-Millis model for the underlying QCP (*19*). In contrast, the naïve expectation for two successive first-order metamagnetic transitions unrelated to the QCP would be lines of qualitatively similar slope and curvature, whereas the boundary of a new phase might have been expected to be dome-shaped.

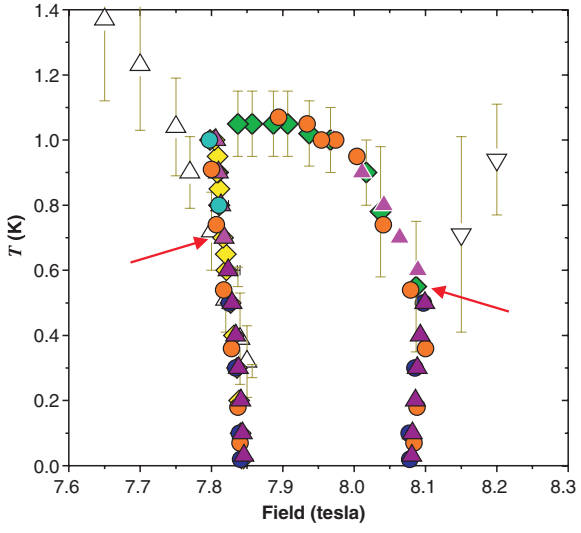


Fig. 3. An empirical phase diagram from collating data from the four measurements contributing to Figs. 1 and 2, combined with the results of thermal expansion measurements. The colored triangles are the loci of peaks in ac susceptibility; the white triangles, the loci of maxima (up) and minima (down); the orange circles are derived from dc magnetization; the yellow diamonds, the loci of maxima in $d\rho/dH$; the green diamonds, the loci of maxima in $d^2\rho/dT^2$; and the blue circles, the loci of maxima in the magnetostriction. The red arrows mark the temperatures of critical points deduced from the temperatures above which no peak in χ'' is observable. We believe that all the data are consistent with the presence of a phase enclosed with first-order phase boundaries for temperatures below these critical points and a second-order line linking them at higher T.

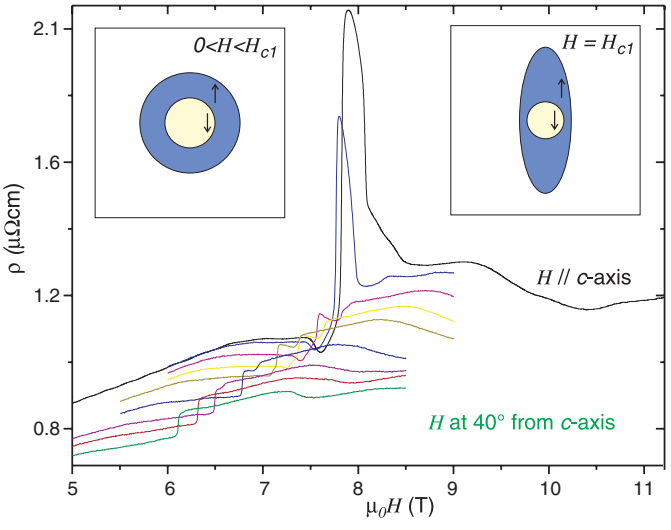


Fig. 4. The resistivity, ρ, at 100 mK as the magnetic field is rotated away from the c axis toward the ab plane at intervals of 5°. The striking peak that we associate with formation of the phase for H // c (black trace) exists only over a narrow range of angle. For angles < 80°, features are still seen, but they are much weaker. (**Insets**) Sketches of the kind of spin-dependent Fermi surface distortion that we discuss in the text. They are not meant to be representative of the true Fermi surfaces but are designed to illustrate the kind of symmetry breaking that we discuss. H_{c1} denotes the field (about 7.8 T) at which the symmetry-broken phase is entered. Below this field and above H_{c2} (the field of about 8.1 T at which the phase is left), both the spin-up and spin-down Fermi surfaces have the same symmetry (left inset). Between H_{c1} and H_{c2}, the spin-up Fermi surface distorts to a lower symmetry. This symmetry-breaking distortion could equally well be rotated by 90°, so domain formation is likely.

The third piece of experimental support for a connection with the quantum critical physics of the underlying metamagnetism comes from an initial study of the angular dependence of ρ. Previous work on more disordered samples has shown that rotating the field from the c axis toward the ab plane raises the characteristic temperature of the underlying critical point, "detuning" the quantum criticality (*13*). It is, therefore, natural to investigate the effect of changing field angle on the formation of the anomalous phase. To do this for all the thermodynamic quantities contributing to Fig. 3 will be a formidable experimental task, but the resistivity data (Fig. 4) show that the anomalous behavior is well bounded in field angle as well as in field and temperature. Each method of tuning is therefore seen to affect the underlying critical point and the appearance of new phase in a qualitatively similar way.

Although itinerant metamagnets have been studied fairly extensively in the past [for example, (*5*, *20*, *21*)], novel phase formation in the vicinity of the metamagnetism has been reported only in URu_2Si_2 (*22*, *23*). The phenomena seen there may have a similar origin to those reported here, but there are important differences in the physics of the two materials. In URu_2Si_2, f electrons play an important role, and both superconductivity and a much-studied "hidden order" phase are observed in zero applied field. Also, the phase formation seems relatively insensitive to the presence of disorder, in contrast to the present case. Although we certainly do not rule out relevance to the fascinating physics that has been discovered in URu_2Si_2, the following discussion will have its basis purely in the observations reported here on $Sr_3Ru_2O_7$.

The key facts that any model for the new phase needs to address are (i) the existence of a large enhancement in the absolute value of ρ accompanied by a suppression of its temperature dependence, (ii) an itinerant rather than a localized picture, as evidenced by the observation of dHvA oscillations at both high and low fields (*24*) and the absence of a large change in the Hall

coefficient as the anomalous phase is entered (*25*), (iii) a magnetization that increases both on entry to and exit from the new phase as a function of field, (iv) a magnetization whose divergence is cut off as the phase is entered from high temperatures at constant field and whose absolute value shows a kink but no sharp decrease as the phase boundary is crossed, and (v) existence of the phase over only a narrow range of applied field angle.

These facts place some fairly tight constaints on theory. For example, they appear to rule out a partially gapped state such as that recently discussed in relation to the hidden order of URu_2Si_2 (*26*). One scenario, however, seems to offer a plausible explanation for all the observations. In a standard picture of itinerant metamagnetism, the system develops the additional moment by a sudden extra exchange polarization of the Fermi surface, changing the relative volumes of the spin-up and spin-down parts (Fig. 4, left inset). The extra exchange splitting at the metamagnetic transition is driven by the Zeeman splitting of the Fermi surface in an applied field, allowing one or both spin species to access a higher density of states than is available in zero field so that the system satisfies the Stoner criterion. Many-body enhancements also play the role of reducing the energy scale of the important features in the density of states [to ~10 K in $Sr_3Ru_2O_7$ (*11*, *12*)].

A natural assumption is that in **k** space, the metamagnetic exchange splitting will respect the fourfold symmetry of the RuO_2 planes. We postulate, however, that something more unusual can happen in $Sr_3Ru_2O_7$. Its quasi-two-dimensional electronic structure will incorporate fourfold density of states maxima because of van Hove singularities in the electronic structure. Indeed, the importance of the Fermi level lying close to these has been discussed in the context of the standard metamagnetism (*27*). We believe that the main features of our observations would be explained by the onset of a spontaneous, twofold, symmetric Fermi-surface distortion in which the Fermi surface of one spin species elongates along the direction of two of the four available density of states peaks (Fig. 4, right inset). This Fermi surface, which breaks the original fourfold symmetry, persists throughout the anomalous new phase until a second phase transition restores the original symmetry by polarizing equally along the orthogonal direction. This scenario has a number of attractive features. It would naturally account for the facts that the magnetization of the system rises at each transition and that it is not diminished on entering the new phase as a function of temperature at constant field. It involves no partial gapping of the Fermi surface but still gives a way to understand the large resistive peak in terms of domain formation between regions in which the symmetry-breaking distortion is rotated by 90°. Alignment of these domains by the inplane field component would also likely be a factor in the rapid disappearance of the peak with field angle.

Symmetry-breaking Fermi-surface distortions have been discussed theoretically in various contexts since the advent of Landau's Fermi liquid theory and the subsequent work of Pomeranchuk (*9*, *28*–*33*). However, there have been, to our knowledge, no conclusive observations in real metals. This raises the question of what is so special about $Sr_3Ru_2O_7$. We believe that part of the answer lies in the underlying metamagnetic QCP. As stressed in (*12*–*14*), the fluctuations associated with an itinerant metamagnetic QCP are rather unusual. They are fluctuations of the Fermi surface itself and so would act to soften the original fourfold Fermi surface, making it easier for it to lower its total energy by adopting a symmetry-breaking distortion. The other important feature of $Sr_3Ru_2O_7$ is the level of purity to which it can be grown. Just like anisotropic superconducting gaps, anisotropic Fermi-surface distortions would be averaged away by sufficiently strong disorder scattering (*34*). The fact that we observe a strong dependence of the new phase on elastic scattering is consistent with this behavior being driven by a Fermi-surface distortion rather than, for example, a structural transition.

We suggest the above scenario as an intuitively appealing possible explanation for the observations that are the core of this paper but do not claim it to be the only way to account for the data. Real-space phase separation or valence transitions are two other possibilities worthy of investigation (*35*). More theoretical and experimental work will be necessary to tell for certain (*36*).

References and Notes

1. S. L. Sondhi, S. M. Girvin, J. P. Carini, D. Shahar, *Rev. Mod. Phys.* **69**, 315 (1997).
2. S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 1999).
3. T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, M. P. A. Fisher, *Science* **303**, 1490 (2004).
4. For example, D. Belitz, T. R. Kirkpatrick, *J. Low Temp. Phys.* **126**, 1107 (2002).
5. G. R. Stewart, *Rev. Mod. Phys.* **73**, 797 (2001).
6. N. D. Mathur *et al.*, *Nature* **394**, 39 (1998).
7. For example, H. L. Stormer, T. R. Kirkpatrick, *Rev. Mod. Phys.* **71**, S298 (1999).
8. S. Ikeda, Y. Maeno, S. Nakatsuji, M. Kosaka, Y. Uwatoko, *Phys. Rev. B* **62**, R6089 (2000).
9. Y. A. Pomeranchuk, *Sov. Phys. JETP* **8**, 361 (1959) [translation from *Zh. Eksp. Teor. Fiz.* **35**, 524 (1959)].
10. A. P. Mackenzie, Y. Maeno, *Rev. Mod. Phys.* **75**, 657 (2003).
11. S. Ikeda, Y. Maeno, S. Nakatsuji, M. Kosaka, Y. Uwatoko, *Phys. Rev. B* **62**, R6089 (2000).
12. R. S. Perry *et al.*, *Phys. Rev. Lett.* **86**, 2661 (2001).
13. S. A. Grigera *et al.*, *Phys. Rev. B* **67**, 214427 (2003).
14. S. A. Grigera *et al.*, *Science* **294**, 329 (2001).
15. R. S. Perry, Y. Maeno (2004); available online at http://arxiv.org/abs/cond-mat/?0403572.
16. R. S. Perry *et al.*, *Phys. Rev. Lett.* **92**, 166602 (2004).
17. The dissipative peak in χ'' is accompanied both by a peak in the imaginary part in field-modulated resistivity measurements and by hysteresis between up and down sweeps in χ' (*25*).
18. Our identification of the second-order phase boundary in Fig. 3 has its basis in the locus of the kinks in the magnetization coinciding with that of maxima in $d^2\rho/dT^2$.
19. A. J. Millis, A. J. Schofield, G. G. Lonzarich, S. A. Grigera, *Phys. Rev. Lett.* **88**, 217204 (2002).
20. T. Goto, K. Fukamichi, H. Yamada, *Physica B* **300**, 167 (2001).
21. J. Flouquet, P. Haen, S. Raymond, D. Aoki, G. Knebel, *Physica B* **319**, 251 (2002).
22. M. Jaime, K. H. Kim, G. Jorge, S. McCall, J. A. Mydosh, *Phys. Rev. Lett.* **89**, 287201 (2002).
23. N. Harrison, M. Jaime, J. A. Mydosh, *Phys. Rev. Lett.* **90**, 96402 (2003).
24. R. A. Borzi *et al.*, *Phys. Rev. Lett.* **92**, 216403 (2004).
25. S. A. Grigera *et al.*, unpublished data.
26. P. Chandra, P. Coleman, J. A. Mydosh, V. Tripathi, *Nature* **417**, 831 (2002).
27. B. Binz, M. Sigrist, *Europhys. Lett.* **65**, 816 (2004).
28. It is, perhaps, debatable whether the Fermi-surface distortion discussed here should be classed as a Pomeranchuk instability. In Pomeranchuk's original work, the presence of a lattice was not necessary for the many-body distortion to take place. Here, it is likely that the lattice does play a role by modulating the **k** dependence of the density of states and possibly also through the strong magnetostructural coupling demonstrated in Fig. 2. This does not, however, mean that many-body effects are unimportant to the observed behavior. In fact, they are probably crucial to what takes place. First, they renormalize the energy scale of the density of states down to a few kelvin, and secondly they dominate the quantum critical fluctuations that we presume to be vital in softening the undistorted Fermi surfaces. In this sense, we believe that it is appropriate to describe the model that we propose with use of the term "spin-dependent Pomeranchuk instability."
29. L. P. Gor'kov, A. Sokol, *Phys. Rev. Lett.* **69**, 2586 (1992).
30. C. J. Halboth, W. Metzner, *Phys. Rev. Lett.* **85**, 5162 (2000).
31. V. Oganesyan, S. A. Kivelson, E. Fradkin, *Phys. Rev. B* **64**, 195109 (2001).
32. I. Khavkine, C.-H. Chung, V. Oganesyan, H.-Y. Kee, (2004); available online at http://arxiv.org/abs/cond-mat/?0402565.
33. H. Y. Kee, Y. B. Kim (2004); available at http://arxiv.org/abs/cond-mat/?0408004+ (partly motivated by the results and discussion presented here).
34. A. G. Green, S. A. Grigera, B. D. Simons, unpublished data.
35. A real-space phase separation giving a coexistence region between 7.8 and 8.1 T cannot be completely ruled out on the basis of our data. However, we believe that it is unlikely. Both hysteretic dissipation (measured by χ'') and hysteresis of the dc magnetization are confined to narrow regions about the first-order phase boundaries identified in Fig. 3. A phase coexistence might be expected to yield maximum dynamic dissipation where the mixture is closest to 50/50, namely in between these two lines.
36. Some experiments that would provide strong supporting evidence of "Pomeranchuck domains" are study of the resistivity under uniaxial inplane pressure anisotropy and study of the inplane resistivity tensor as a function of the polar angle in tilted magnetic fields. Further insight might be gained by using spatially resolved probes such as scanning tunnelling spectroscopy. All these experiments are planned.

37. We are pleased to acknowledge useful discussions with K. Ishida, G. G. Lonzarich, V. Oganesyan, and B. D. Simons. This work was supported by Engineering and Physical Sciences Research Council, The Leverhulme Trust, and the Royal Society (UK), and by the Japanese Society for the Promotion of Science and the Ministry of Education, Culture, Sports, Science, and Technology (Japan).

20 August 2004; accepted 5 October 2004

The Roaming Atom: Straying from the Reaction Path in Formaldehyde Decomposition

D. Townsend,[1,2] S. A. Lahankar,[3] S. K. Lee,[1,2,3] S. D. Chambreau,[3] A. G. Suits,[1,2,3]* X. Zhang,[4] J. Rheinecker,[4] L. B. Harding,[5] J. M. Bowman[4]*

We present a combined experimental and theoretical investigation of formaldehyde (H_2CO) dissociation to H_2 and CO at energies just above the threshold for competing H elimination. High-resolution state-resolved imaging measurements of the CO velocity distributions reveal two dissociation pathways. The first proceeds through a well-established transition state to produce rotationally excited CO and vibrationally cold H_2. The second dissociation pathway yields rotationally cold CO in conjunction with highly vibrationally excited H_2. Quasiclassical trajectory calculations performed on a global potential energy surface for H_2CO suggest that this second channel represents an intramolecular hydrogen abstraction mechanism: One hydrogen atom explores large regions of the potential energy surface before bonding with the second H atom, bypassing the saddle point entirely.

As chemical kinetic theory has evolved from Eyring's work in the 1930s through the Rice-Ramsperger-Kassel-Marcus (RRKM) and related statistical approaches, the transition-state concept has remained paramount. Reactants are assumed to proceed along the lowest energy pathway to products, with the configuration at the energetic maximum of this pathway termed the transition state. The properties of the transition state, in particular its geometry and vibrational frequencies, have profound importance for determining the rate of reaction, so studies of reaction rate theory have focused mainly on elucidating these properties. Absent from this paradigm is the possibility of two or more distinct pathways to the same products. We present here a combined theoretical and experimental study of formaldehyde (H_2CO) dissociation that supports two distinct active pathways to H_2 and CO products. Whereas the dominant pathway involves a conventional transition state, the alternative pathway is an intramolecular hydrogen abstraction that avoids the transition state region entirely. These results, considered with recent work on the $CH_3 + O$ (*1*) and $O + C_2H_6$ (*2*) reactions and other cases in which regions of deep potential energy wells are avoided in reaction (*3*), suggest that transition-state theories are incomplete in their description of chemical reactivity.

The study of unimolecular dissociation of vibrationally excited molecules in the ground electronic state has been central to the development of transition-state theories, and the nature of the transition state itself has been the focus of many of these studies (*4*, *5*). Formaldehyde has literally become a textbook example (*6*) with which to study these issues. It is one of the simplest molecules in which to examine the correlation between rotational and vibrational excitation in the products. Moreover, laser excitation efficiently prepares H_2CO on the ground electronic state with specific amounts of internal energy.

Moore and co-workers have explored the unimolecular reaction dynamics of formaldehyde in great depth (*7*–*12*). In studies of both H_2 and CO product state distributions and state-resolved H_2 Doppler profiles, they found about 65% of the available energy is released in translation; the rest of the energy is partitioned between strong CO rotational excitation and modest H_2 vibrational excitation. They found little energy in CO vibrational excitation or in rotational excitation of

[1]Department of Chemistry, Stony Brook University, Stony Brook, NY 11794, USA. [2]Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, USA. [3]Department of Chemistry, Wayne State University, Detroit, MI 48202, USA. [4]Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, GA 30322, USA. [5]Chemistry Division, Argonne National Laboratory, Argonne, IL 60439, USA.

*To whom correspondence should be addressed. E-mail: asuits@wayne.edu; jmbowma@emory.edu

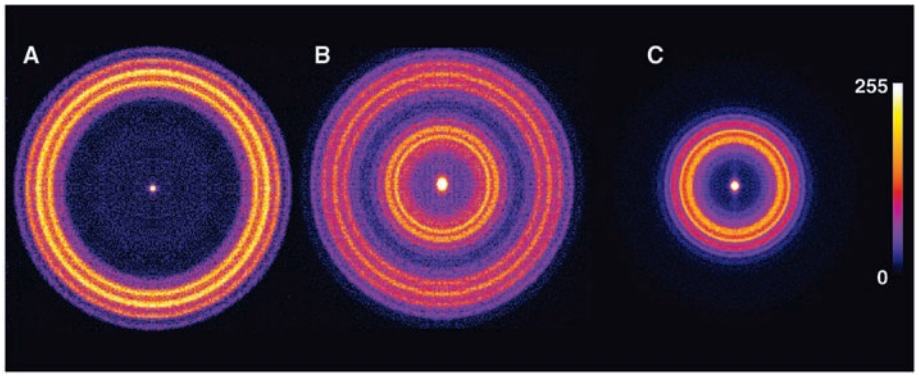


Fig. 1. DC sliced images of CO ($v = 0$) after dissociation of H_2CO at 30,340.1 cm^{-1} excess energy for the CO product rotational levels $j_{CO} = 40$ (**A**), $j_{CO} = 28$ (**B**), and $j_{CO} = 15$ (**C**).

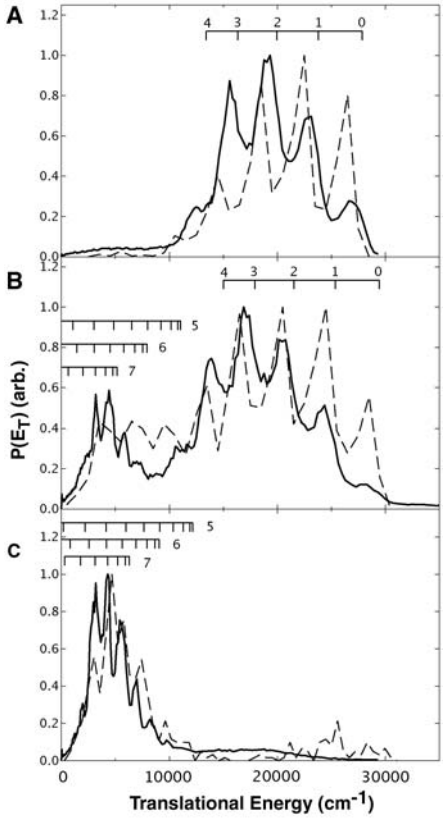


Fig. 2. (**A** to **C**) Solid lines are experimental translational energy distributions obtained from the corresponding images in Fig. 1. Dashed lines are translational energy distributions obtained from the trajectory calculations. Markers indicate positions for correlated H_2 (v and $j = 1$) vibrational levels for v from 0 to 4 and odd rotational levels for H_2 v from 5 to 7. $P(E_T)$ is the probability of a given translational energy (E_T) (arbitrary units).

H_2. All of these observations could be accounted for qualitatively by a skewed transition-state structure in which both hydrogen atoms are on the same side of the CO molecule. These results have also been accounted for semiquantitatively by quasiclassical trajectory calculations which were initiated at this transition state (*13*–*16*).

However, one subtle feature of the measurements did not fit well with this conventional understanding of the dynamics. At energies above 30,300 cm^{-1}, H_2CO can undergo an alternative dissociation to H + HCO (the radical channel). In a prescient paper, van Zee *et al*. noted (*12*) that, at energies below this threshold, the CO rotational state (j_{CO}) distributions were simply Gaussian-shaped, peaking near $j_{CO} = 45$, and were well understood to arise from energy and angular momentum conservation given the nature of the transition state and the exit channel interactions. However, above the threshold for the H + HCO dissociation channel, the CO rotational distributions exhibited a shoulder toward lower rotational levels. van Zee and co-workers proposed two possible explanations for the low-j_{CO} component. One explanation was that at higher energies the transition-state region may sample additional geometries that lead to reduced exit impact parameters, hence lower rotational excitation in the CO and higher overall translational and vibrational energy. The second possibility was that the low j_{CO} was related to the opening of the radical channel and thus a distinct new pathway to formation of molecular products via intramolecular hydrogen abstraction.

To resolve this issue, we performed high-resolution direct current (DC) slice imaging (*17*) measurements of specific CO rovibrational levels after dissociation of formaldehyde above the threshold for the radical channel. These images provide the correlated H_2 vibrational distributions with vibrational and partial rotational resolution. The experimental measurements are combined with quasiclassical trajectory calculations performed on a new, high-level, global potential energy surface for H_2CO (*18*).

The ion imaging experiment (*19*, *20*) involves laser excitation of formaldehyde in a molecular beam under collisionless conditions to the first electronically excited singlet state at 30,340.1 cm^{-1} (*21*), one of several dissociation energies used by van Zee *et al*. This excitation includes one quantum in the CO stretch and three in the out-of-plane bend. Internal conversion then produces a distribution of vibrationally excited ground-state molecules that dissociate some picoseconds later (*22*). The product, CO, is probed in the dominant ground vibrational state on specific rotational levels by using DC slice imaging. By using weak DC electric focusing fields and resonant multiphoton ionization, we are able to image the central slice through the recoiling product spheres. This yields high-resolution state-resolved velocity distributions (Fig. 1). These images represent the velocity distributions for specific indicated quantum states of the CO product, and the structure in the images reflects the internal energy distribution of the H_2 co-product for the particular CO level being probed because the total energy is fixed by the photolysis laser.

In the image for $j_{CO} = 40$ (Fig. 1A), the rings represent formation of the H_2 co-product in vibrational states (v) from 0 to 3 with substantial translational energy release. Integration of the data in the images and conversion from velocity to total energy yields the translational energy distribution (Fig. 2A, solid line). This resulting H_2 vibrational distribution peaks at $v = 2$ with the bulk of the available energy appearing in translation, consistent with Moore's measurements. The image obtained for $j_{CO} = 28$ (Fig. 1B) reveals a bimodal internal energy distribution in the H_2 co-product, strongly suggesting a distinct dissociation mechanism. This rotational level is intermediate between the peak of the rotational distribution and the low-j_{CO} shoulder reported by van Zee *et al*. The corresponding translational energy distribution (Fig. 2B), showing partial rotational resolution of the correlated H_2, reveals that the slower CO product is formed in conjunction with H_2 in vibrational levels up to $v = 7$. Lastly, the image for $j_{CO} = 15$ (Fig. 1C) in the region of the low-j_{CO} shoulder reported by van Zee *et al*. shows only the slow component, and the translational energy distribution (Fig. 2C) is

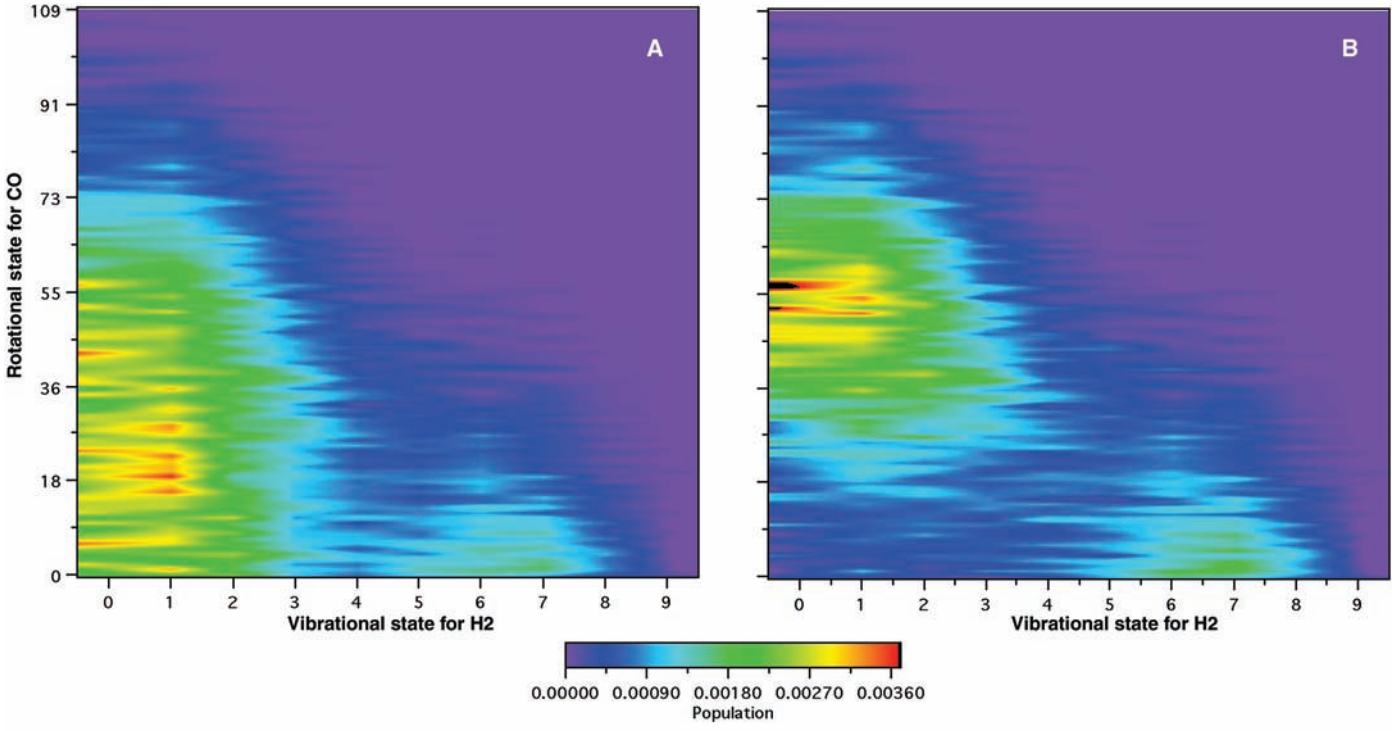


Fig. 3. State correlations obtained from the trajectory calculations: (**A**) CO (all v and j) compared with H_2 (v) and (**B**) CO ($v = 0$ and j) compared with H_2 (v).

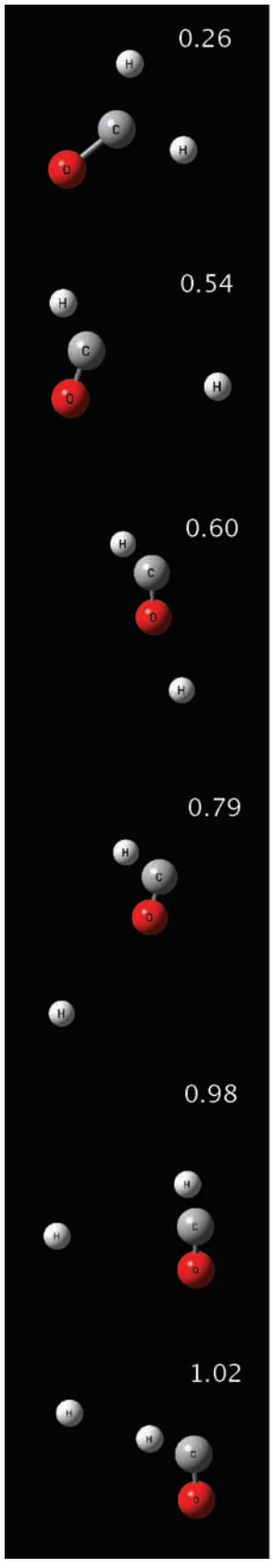


Fig. 4. Six frames of a sample trajectory leading to CO (j_{CO} = 7) and H_2 (v_{H2} = 6) at the indicated times after initiation.

quite similar to the slow part of the distribution in Fig. 2B: In this case, we see only formation of highly internally excited H_2 up to v = 7 and rotational state j = 13.

The accompanying trajectory calculations were performed on a new H_2CO potential energy surface (*18*) constructed from least squares fits to roughly 60,000 high-level ab initio points, followed by a smooth joining of these fits. The surface contains the molecular (H_2 + CO) and radical (H + HCO) dissociation channels as well as the cis and trans isomers of HCOH. The relevant energetics of this surface, including corrections for harmonic zero-point energy, are 28,633 cm^{-1} for the barrier to the molecular channel and 30,325 cm^{-1} for the radical channel, in very good agreement with experiment.

Quasi-classical trajectories on this surface were calculated at a total energy corresponding to the experiment based on a harmonic estimate of the zero-point energy of 5844 cm^{-1}. Two sets of trajectories were performed: In one the initial molecular geometry was set at the formaldehyde minimum, and in the other one CH bond was initially stretched to a distance of 2.1 Å, about twice the equilibrium value (*23*). Roughly 40,000 trajectories were performed for each initial geometry with zero total angular momentum. The generation of initial conditions for these trajectories followed the standard method (*13–16*), and the usual histogram binning procedure was used to obtain final vibration/rotation (v/j) distributions of the H_2 and CO products. The results from the two sets of calculations are identical within the statistical uncertainties of each, confirming the robustness of the results. The translational energy distributions obtained from the calculations (Fig. 2, dashed lines) are in near-quantitative agreement with the experimental data for each CO rotational level. The calculations also yield probabilities for dissociation to given levels of CO rotational and H_2 vibrational excitation, both summed over CO vibrational levels (Fig. 3A) and restricted to v_{CO} = 0 (Fig. 3B). In both cases, there is a bimodal distribution in v_{H2}, with a major peak at v_{H2} = 1 and a smaller peak at v_{H2} = 6. Further in accord with experiment, the vibrationally excited H_2 correlates with a CO rotational distribution peaked near j_{CO} = 10; the vibrationally cold H_2 is formed with rotationally excited CO (peaked near j_{CO} = 45). Calculations have also been done at five energies spanning a 3000-cm^{-1} energy range centered at the energy of the experiment. At the lowest energy, the low-j_{CO} shoulder in the CO rotational distribution is near zero relative to the peak value (j_{CO} = 42), and this contribution grows monotonically as the energy increases. Correspondingly, the second peak in the bimodal v_{H2} distribution is nearly zero relative to the first peak at low v_{H2}. The branching ratio for the molecular channel decreases monotonically from roughly 1.0 to 0.2 (*24*). Thus, whereas the high-v H_2 and low-j_{CO} component of the molecular channel increases with energy, the branching ratio of the molecular products decreases. These results are consistent with the dissociation energy dependence of the low-j_{CO} shoulder reported by van Zee *et al.*

One compelling feature of these calculations is that trajectories may be examined directly for insight into the associated dynamics. A typical trajectory leading to j_{CO} = 41 and v_{H2} = 1 (movie S1) proceeds via the skewed transition state, and the corresponding reaction path is closely followed. In contrast, a trajectory leading to j_{CO} = 7 and v_{HH} = 6 (movie S2) involves one H atom nearly detaching via the H + HCO channel, although lacking sufficient energy for complete dissociation. The H atom meanders in the broad attractive space of the H + HCO surface, far from the reaction path, until it abstracts the other hydrogen atom to yield highly vibrationally excited H_2 and rotationally cold CO. Figure 4 shows six frames of movie S2, illustrating the "roaming" mechanism.

These two trajectories illustrate the dramatically distinct mechanisms underlying the bimodal correlated state distribution seen in the experimental images and trajectory calculations. In one case, we have what may be described as the conventional dissociation pathway involving the well-characterized transition state. For this dominant mechanism, transition-state theories are appropriate and may be used to predict the dissociation rates and product state distributions. For the second case, we find compelling evidence in support of van Zee's conjecture invoking the H + HCO channel. This pathway is analogous to recently calculated trajectories for the O + CH_3 reaction (*1*) involving near H elimination followed by intramolecular H abstraction, which accounted for CO yields seen experimentally.

We suspect this roaming atom mechanism is a common pathway. A key question remaining is whether diatomic products besides H_2, or even polyatomic products, may be formed by such a mechanism. It may be more common for the roaming species to be a hydrogen atom that can rapidly explore the accessible regions of the surface. Perhaps, however, an atom or group besides H could be the abstraction target.

References and Notes

1. T. P. Marcy *et al.*, *J. Phys. Chem. A* **105**, 8361 (2001).
2. T. Yan, C. Doubleday, W. L. Hase, *J. Phys. Chem. A*, in press (ASAP Web release 16 July 2004, DOI 10.1021/jp048150).
3. L. Sun, K. Song, W. L. Hase, *Science* **296**, 875 (2002).
4. T. Baer, W. L. Hase, *Unimolecular Reaction Dynamics: Theory and Experiment* (Oxford Univ. Press, Oxford, 1996).

5. C. B. Moore, *Faraday Discuss.* **102**, 1 (1995).
6. R. Schinke, *Photodissociation Dynamics* (Cambridge Univ. Press, Cambridge, 1993).
7. P. Ho, D. J. Bamford, R. J. Buss, Y. T. Lee, C. B. Moore, *J. Chem. Phys.* **76**, 3630 (1982).
8. D. J. Bamford, S. V. Filseth, M. F. Foltz, J. W. Hepburn, C. B. Moore, *J. Chem. Phys.* **82**, 3032 (1985).
9. D. Debarre *et al.*, *J. Chem. Phys.* **83**, 4476 (1985).
10. T. J. Butenhoff, K. L. Carleton, M. C. Chuang, C. B. Moore, *J. Chem. Soc. Faraday Trans.* **85**, 1155 (1989).
11. T. J. Butenhoff, K. L. Carleton, C. B. Moore, *J. Chem. Phys.* **92**, 377 (1990).
12. R. D. van Zee, M. F. Foltz, C. B. Moore, *J. Chem. Phys.* **99**, 1664 (1993).
13. Y.-T. Chang, C. Minichino, W. H. Miller, *J. Chem. Phys.* **96**, 4341 (1992).
14. X. Li, J. M. Millam, H. B. Schlegel, *J. Chem. Phys.* **113** (2000).
15. W. Chen, W. L. Hase, H. B. Schlegel, *Chem. Phys. Lett.* **228**, 436 (1994).
16. J. L. Rheinecker, X. Zhang, J. M. Bowman, in preparation.
17. D. Townsend, M. P. Minitti, A. G. Suits, *Rev. Sci. Instruments* **74**, 2530 (2003).
18. X. Zhang, S. Zou, L. B. Harding, J. M. Bowman, *J. Phys. Chem. A* **108**, 8980 (2004).
19. D. W. Chandler, P. L. Houston, *J. Chem. Phys.* **87**, 1445 (1987).
20. A. Eppink, D. H. Parker, *Rev. Sci. Instrum.* **68**, 3477 (1997).

21. The experiment features a pulsed molecular beam of formaldehyde directed into a vacuum chamber along the axis of an imaging time-of-flight mass spectrometer. The beam is crossed at right angles by counter-propagating photolysis and probe lasers. The photolysis laser is the frequency-doubled output of a yttrium-aluminum-garnet–Nd (Nd:YAG) pumped dye laser tunable around 330 nm. The probe laser is the frequency-tripled output of a second Nd:YAG pumped dye laser, tunable around 230 nm. The probe uses the well-known 2-plus-1 resonance-enhanced ionization of CO via the (B-X) Q-branch transition, affording, in this case, detection of single CO rotational levels for $j = 10$ and higher. Data images are recorded with centroiding without filtering or smoothing and then fourfold symmetrized to improve signal to noise.

22. C. B. Moore, J. C. Weisshaar, *Annu. Rev. Phys. Chem.* **34**, 525 (1983).

23. A normal-mode analysis was done at these two initial configurations, and normal modes associated with CH motion were clearly identified. The sum of harmonic zero point energies for all other modes was subtracted at the outset. The excess energy was then assigned as kinetic energy to the CH normal modes. For an initial configuration at the H_2CO minimum, this excess energy was equally partitioned between the symmetric and antisymmetric CH stretches. For a stretched H_2CO initial configuration, the excess energy was given to the one real-frequency CH-stretch normal mode. To obtain good statistics for the quasiclassical distributions, we collected those results over a j_{CO} range of $\pm$ 2. Also, the translational energy range was discretized in bins of 1000 cm^{-1} width except for the case $j_{CO} = 15$, where the width was 500 cm^{-1}. Variation of these bin widths by $\pm$ 20% had no significant effect on the plotted distributions.
24. In quasi-classical trajectory calculations, the product molecules can be formed with less than zero-point energy, which violates a strict quantum condition. In the present calculations, we see this effect by observing a small amount of H + HCO products below the strict threshold energy of 30,300 cm^{-1}. Thus, the branching ratio to form H_2 + CO is not exact unity as it should be; however, it is near unity. One difficulty in establishing a more reliable value at these lower energies is that a significant number of trajectories became very long-lived and had to be terminated before dissociation occurred.

25. This work was supported by the director, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, U.S. Department of Energy, under contracts DE-AC02-9810886 and DE-FG02-04ER15593 (A.G.S.), DE-FG02-97ER14782 (J.M.B.), and W-31-109-ENG-38 (L.B.H.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1104386/DC1
Movies S1 and S2

23 August 2004; accepted 7 October 2004
Published online 21 October 2004;
10.1126/science.1104386
Include this information when citing this paper.

Sea Urchin Spine Calcite Forms via a Transient Amorphous Calcium Carbonate Phase

Yael Politi,[1] Talmon Arad,[2] Eugenia Klein,[2] Steve Weiner,[1] Lia Addadi[1]*

The skeletons of adult echinoderms comprise large single crystals of calcite with smooth convoluted fenestrated morphologies, raising many questions about how they form. By using water etching, infrared spectroscopy, electron diffraction, and environmental scanning electron microscopy, we show that sea urchin spine regeneration proceeds via the initial deposition of amorphous calcium carbonate. Because most echinoderms produce the same type of skeletal material, they probably all use this same mechanism. Deposition of transient amorphous phases as a strategy for producing single crystals with complex morphology may have interesting implications for the development of sophisticated materials.

Many organisms—including mollusks, echinoderms, calcisponges, corals, certain algae, and others—form their hard parts out of calcium carbonate minerals (*1*, *2*). The echinoderms, which include sea urchins, sea stars, and brittle stars, among others, are among the few groups that form skeletal parts consisting of large single crystals of Mg-bearing calcite, some of which are several centimeters long (*3*, *4*). These single crystals usually have smooth, continuously curved surfaces that form a three-dimensional fenestrated mineral network surrounding micrometer-sized spaces that are occupied by living cellular tissue. This so-called stereom appears to be an "all purpose material," used by members of the whole phylum (*5*). It is molded into different shapes and sizes and is thus adapted to fulfill different functions.

The strategy evolved by echinoderms to build their spines, test plates, and ossicles has been the subject of study for more than a hundred years (*6*, *7*). The crystals form inside a syncytium (*8*), a membrane envelope produced by many cells. The cells provide the raw materials necessary for constructing the growing single crystals (*9*, *10*). A key question is how a single crystal of calcite can be molded into such convoluted shapes, when inorganically formed calcite crystals are rhombohedra with flat crystal faces. Another perplexing issue is how the cells can efficiently provide the ions and, at the same time, efficiently remove water, during and after crystal deposition. This question becomes critical if it is assumed that the crystal grows out of an aqueous solution saturated with calcium carbonate (*11*).

Insights into these questions have been gained from studies of calcitic spicule formation in sea urchin larvae (*12*). The larval spicules grow on a single calcite crystal seed by transformation of a transient amorphous calcium carbonate (ACC) phase (*13*). ACC is apparently fed into the syncytium by cells in the form of ACC-containing vesicles (*14*). Thus, packages of ACC are delivered to the crystal deposition site and then transform in a controlled manner into calcite single crystals. There is also no discernible aqueous phase around the growing spicule (*14*). Mollusk larvae also form their shells via an ACC precursor phase, which then transforms into aragonite (*15*). Because this strategy may be a unique feature of larvae adapted to a swimming mode of life in the oceans, it was important to determine whether or not adult echinoderms or mollusks also use transient ACC for forming their hard parts. We show that adult sea urchins do build their regenerating spines via the deposition of an amorphous ACC precursor phase.

Sea urchins are able to regenerate spines that break. Spine regeneration first involves a wound-healing phase during which the epidermis is reconstituted around the broken spine (*16*, *17*). Within this space, the skeleton-

[1]Department of Structural Biology, [2]Chemical Research Support Unit, Weizmann Institute of Science, 76100 Rehovot, Israel.

*To whom correspondence should be addressed. E-mail: lia.addadi@weizmann.ac.il

forming sklerocytes build a new syncytium, which is in contact with the stump of the old spine. The new crystal is presumably nucleated epitaxially on the old spine (*18*), because the regenerated and old spines together diffract x-rays as one single crystal. The regeneration process is thus considered to be very similar to the original spine growth. Regeneration starts in the form of small projections emerging from the broken stereom surface. These microspines initially grow parallel to the mature spine long axis, which is also the *c* crystallographic axis of calcite, and then form the complex three-dimensional fenestrated structure (*19*) (Fig. 1, A and B). As the process is initiated with the break, a real-time investigation of rapid skeleton deposition is possible. We used etching in water, infrared spectroscopy, electron diffraction, and environmental scanning electron microscopy (ESEM) to investigate the possible involvement of ACC in spine regeneration of the sea urchin *Paracentrotus lividus*.

The solubility of ACC (200 mg/liter) is 30 times greater than the solubility of calcite (6.7 mg/liter) (*20*, *21*). Etching in water thus reveals the presence of ACC, which is selectively dissolved out of a mixture of ACC and calcite. Four-day-old regenerated microspines were etched in water immediately after severing from the spine (*22*). Extensive etching occurred in a thin (100 to 200 nm thick) surface layer (Fig. 1D). These etched surfaces were compared to regenerated microspines that were left at room temperature in air for 1 month and then etched. The latter were not affected by the treatment (Fig. 1E) and closely resembled the untreated spines (Fig. 1C).

Fourier transform infrared (FTIR) was used to further investigate the nature of the thin surface layer (Fig. 2, A and C). The characteristic broadening of the in-plane carbonate bending peak at 713 cm^{-1}, relative to the out-of-plane bending peak at 876 cm^{-1}, is well documented for ACC-containing sea urchin larval spicules (*13*). Regenerated spines, cleaned of organic matrix, were dried in ethanol and sonicated into acetone after rapid freeze shock treatment in liquid nitrogen. Sonication detached small particles from the surface of the microspines. The infrared spectrum of these particles shows a sharp peak at 876 cm^{-1}, whereas the 713 cm^{-1} peak comprises a sharp peak superimposed on a very broad peak (Fig. 2A, inset). The spectrum of crystalline calcite taken from the mature part of the same spines (Fig. 2B) was then subtracted from the latter. The residual spectrum still shows a sharp out-of-plane bending peak at 876 cm^{-1}, but the 713 cm^{-1} peak consists only of a broad hump (Fig. 2C), clearly indicative of the presence of ACC. A broad peak at ~3500 cm^{-1} is also enhanced in the subtracted spectrum, relative to calcite. This may be due to water molecules that stabilize the amorphous phase and are intimately associated with it. The presence of some water also in the mature crystalline part of the spine indicates, however, that this may be interstitial water or water associated with residual organic matrix.

Particles removed from regenerating spines by the same procedure as described above were dispersed on an electron microscope grid and examined by electron imaging and diffraction (Fig. 3) (*22*). The particles are ~50 nm in size. Particles obtained from fresh regenerated spines did not produce an electron diffraction pattern (Fig. 3A). The same particles, examined after 3 weeks on the grid and kept in air, were crystalline and produced a diffraction pattern typical of calcite (Fig. 3B).

Etching, IR spectroscopy, and transmission electron microscopy thus provide independent evidence that the regenerated microspines are covered by a thin layer of ACC that, with time, transforms into crystalline calcite. The crystallization process occurs both in situ and in particles removed from the spine.

We were also able to directly observe the crystallization of the amorphous sheath and tip of the microspines into calcite in a wet atmosphere, using ESEM (Fig. 4) (*22*). The electron beam provides the energy required to induce crystallization. The tip of the microspine first undergoes a reduction in volume (Fig. 4A; compare with Fig. 4, B

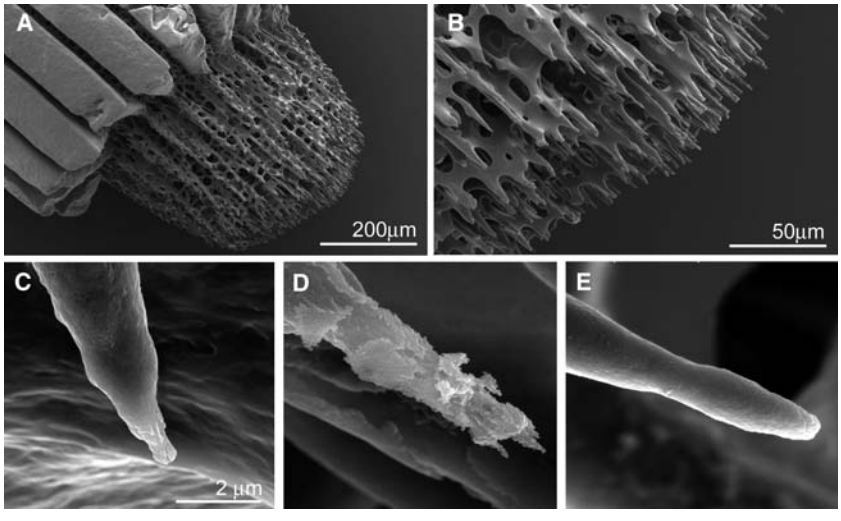


Fig. 1. Scanning electron micrographs of regenerating spines. (**A**) Five-day-old regenerated spine growing on the original broken spine. (**B**) Higher magnification view of the tip of the new growth, showing the typical stereom structure and the protruding newly formed microspines. (**C**) One microspine formed after 4 days of regeneration, observed fresh. (**D**) Four-day-old microspine, etched in water while fresh. (**E**) Four-day-old microspine, etched in water 1 month after regeneration. All the spines were observed after removal of the exposed organic material by a 3% NaOCl solution. Etching was performed by immersing cleaned spines in double distilled water for 12 hours.

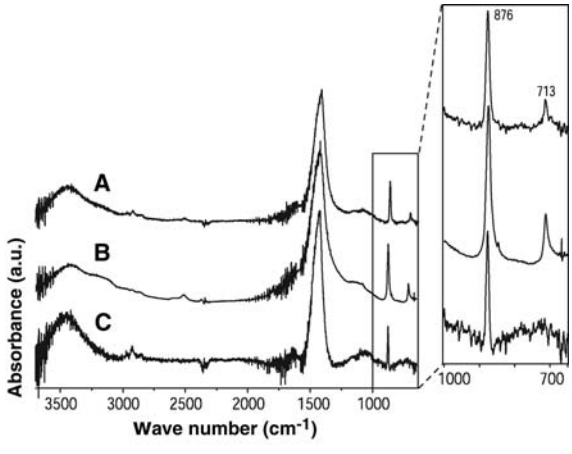


Fig. 2. FTIR spectra of particles removed from fresh regenerated spines by freeze shock and sonication. The intensities are normalized to the carbonate stretching peak at 1420 cm^{-1}. (**A**) IR spectrum of the freshly removed particles. The spectrum corresponds to that of a mixture of ACC and calcite (ratio of intensities I_{876}/I_{713} = 4.25). (**B**) IR spectrum of material removed from the mature part of the spines. The spectrum corresponds to that of crystalline calcite (ratio of intensities I_{876}/I_{713} = 3). (**C**) IR spectrum of the particles in (A), after subtraction of the spectrum of crystalline calcite from the old part of the spine [taken from (B)]. The spectrum corresponds to that of ACC.

and C). A single crystal is then observed to form as thin plates with well-defined edges (Fig. 4, D and E). The delimiting faces subtend, between them and with the axis of the microspine, dihedral angles matching those expected for the (001), (102), and (010) faces of a calcite crystal oriented with the *c* axis parallel to the long axis of the spine. The crystal then grows at the expense of the amorphous material that is consumed in the electron-irradiated region (Fig. 4, F and G), whereas the nonirradiated region remains unmodified (Fig. 4, C and H). When the same procedure is performed on old regenerated spines, or on an old part of the same spine, no such changes are observed, and the spine preserves its dimensions and appearance.

The reduction in spine volume upon crystallization may indicate expulsion of water, which would be in agreement with the IR data (Fig. 2). This is interesting given that stable ACC contains 15 weight % water, corresponding to the molar composition $CaCO_3 \cdot H_2O$ (*23*). In contrast, transient ACC in sea urchin larval spicules contains little or no water (*24*). If the same mechanism is also valid for the growth of adult spines, this would imply that calcium carbonate is first deposited as hydrated ACC, and then dehydrates prior to or concomitant with crystallization. The shrinkage might, however, also be associated with reorganization of the material into a crystalline lattice with or without concomitant expulsion of interstitial water. Irrespective of the mechanism of ACC-to-calcite transition, the introduction and deposition of calcium carbonate as ACC may well provide some answers to the perplexing questions raised above; namely, how the spine is shaped, how the ions are efficiently introduced, and how water is expelled. ACC is introduced into the syncytium as an isotropic noncrystalline solid and can thus be molded into any shape. The solid itself is, of course, a very concentrated source of ions. Water has to be expelled in only small (at most 15%) amounts. This is orders of magnitude lower than if the crystal were deposited directly from solution. The subsequent transformation of the amorphous phase into a composite crystalline solid with much better mechanical properties (*25*) leads to a functional skeleton. A similar strategy could be used to form synthetic materials with minimal porosity and maximal shape flexibility and has been explored in vitro (*26–29*).

Our investigation of the regeneration of the adult sea urchin spine shows that the skeletal hard part forms through an amorphous precursor phase. Most members of this phylum form the same type of skeletal material (*2*), the stereom, with similar properties: large single crystals of magnesium-containing calcite, which break with conchoidal fracture and are reinforced by intracrystalline organic matrix. We therefore suggest that all echinoderms probably use this mineral-formation process. It is conceivable that many other animal phyla also use the same strategy. Isolated examples have been demonstrated involving, other than ACC, amorphous ferric oxide and amorphous calcium phosphate [table 3.2 in (*2*)]. With the discovery that this process is used by adult echinoderms, the possibility that it is a widespread strategy seems likely. Furthermore, this strategy of molding macroscopic elements into any desired morphology and endowing them with the strength and the degree of perfection of a single crystal, but with the reduced brittleness and mechanical properties of a glass, may have interesting implications for the development of sophisticated materials.

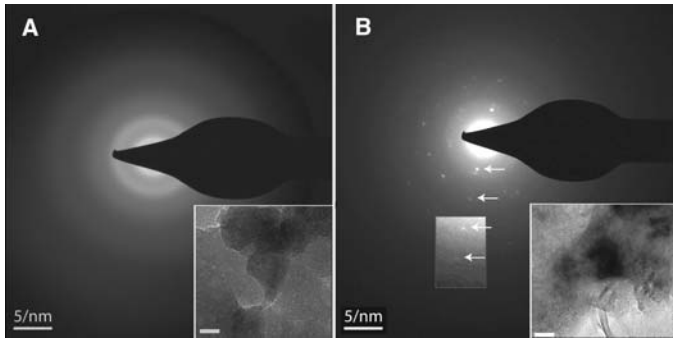


Fig. 3. Transmission electron micrographs (TEM) and electron diffraction patterns of particles removed from fresh regenerated spines. (**A**) Electron diffraction pattern of the particles observed immediately after removal from the spine; the diffuse rings indicate the presence of an amorphous material. (Inset) TEM of the particles. Bar, 50 nm. (**B**) Electron diffraction pattern of the same particles as in (A) after 3 weeks on the grid. The most prevalent diffraction peaks correspond to a *d*-spacing of 3.04 Å, characteristic of the calcite plane {104}. They are detected up to fourth order (arrows, enhanced contrast in window). (Inset) TEM of the particles. Bar, 50 nm. Fresh regenerated spines were subjected to freeze shock and sonication. The separated material was mounted and examined on a marked TEM grid, and the position of the particles was recorded. The same particles were examined again after 3 weeks.

References and Notes

1. K. Simkiss, K. Wilbur, *Biomineralization. Cell Biology and Mineral Deposition* (Academic Press, San Diego, 1989).
2. H. A. Lowenstam, S. Weiner, *On Biomineralization* (Oxford Univ. Press, New York, 1989).

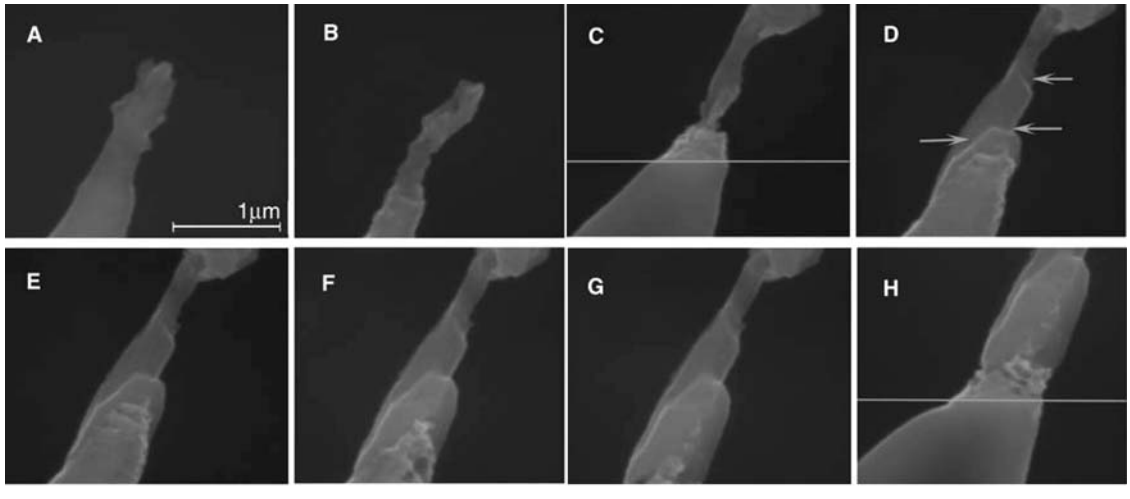


Fig. 4. Environmental scanning electron micrographs (ESEM) of a fresh, newly formed microspine taken in succession over a period of ~15 min. (**A**) Microspine at the beginning of the observation. (**B** to **H**) Pictures taken in succession every 2 to 5 min at the same magnification. Between (B) and (C) and between (G) and (H), the viewing frame was moved along the axis of the microspine to compare the condition of a nonirradiated region of the microspine relative to the irradiated region. Lines in (C) and (H) show the boundaries of the previous frames. Arrows in (D) show the growing crystal faces. Regenerated fresh spines were cleaned from epidermal tissue by NaOCl treatment, washed with water, and immediately transferred wet to an ESEM stub, where they were observed under a humid atmosphere (p_{H_2O} = 6.6 torr, 5°C, 10 kV).

3. K. Towe, *Science* **157**, 1048 (1967).
4. D. M. Raup, *J. Geol.* **67**, 661 (1959).
5. S. Weiner, L. Addadi, H. Wagner, *Mater. Sci. Eng.* **11**, 1 (2000).
6. S. Loven, *Sven. Vetensk. Akad. Handl.* **18**, 1 (1892).
7. W. B. Carpenter, *Brit. Assoc. Adv. Sci. London Rep.* **17**, 93 (1847).
8. K. Okazaki, *Embryologia (Nagoya)* **5**, 283 (1960).
9. J. B. Pilkington, *J. Mar. Biol. Assoc. UK* **49**, 857 (1969).
10. K. Märkel, U. Röser, *Zoomorphology* **103**, 25 (1983).
11. K. Märkel, U. Röser, M. Stauber, *Zoomorphology* **109**, 79 (1989).
12. F. H. Wilt, *Zool. Sci.* **19**, 253 (2002).
13. E. Beniash, J. Aizenberg, L. Addadi, S. Weiner, *Proc. R. Soc. London B Biol. Sci.* **264**, 461 (1997).
14. E. Beniash, L. Addadi, S. Weiner, *J. Struct. Biol.* **125**, 50 (1999).
15. I. M. Weiss, N. Tuross, L. Addadi, S. Weiner, *J. Exp. Zool.* **293**, 478 (2002).
16. K. Märkel, U. Röser, *Zoomorphology* **103**, 43 (1983).
17. B. M. Heatfield, *J. Exp. Zool.* **178**, 233 (1971).
18. P. Dubois, C. P. Chen, *Echinoderm Stud.* **3**, 109 (1989).
19. P. Dubois, L. Ameye, *Microsc. Res. Tech.* **55**, 427 (2001).
20. L. Brecevic, N. A. *J. Cryst. Growth* **98**, 504 (1989).
21. F. Lippmann, *Sedimentary Carbonate Minerals* (Springer, Berlin, 1973).
22. Materials and methods are available as supporting material on *Science* Online.
23. L. Addadi, S. Raz, S. Weiner, *Adv. Mat.* **15**, 959 (2003).
24. S. Raz, P. C. Hamilton, F. H. Wilt, S. Weiner, L. Addadi, *Adv. Funct. Mater.* **13**, 480 (2003).
25. A. Berman, L. Addadi, S. Weiner, *Nature* **331**, 546 (1988).
26. G. F. Xu, N. Yao, I. A. Aksay, J. T. Groves, *J. Am. Chem. Soc.* **120**, 11977 (1998).
27. J. Aizenberg, J. L. Grazul, D. A. Muller, D. R. Hamann, *Science* **299**, 1205 (2003).
28. L. B. Gower, D. J. Odom, *J. Cryst. Growth* **4**, 719 (2000).
29. E. Loste, F. C. Meldrum, *Chem. Commun.*, 901 (2001).

30. We thank M. Cohen for help with the ESEM measurements. We also thank B. Scharfstein, I. Ivry, and Y. Raz (SeaOr Marine Enterprises Ltd.) for providing the animals and facilities for spine regeneration. L.A. is the incumbent of the Dorothy and Patrick Gorman Professorial Chair of Biological Ultrastructure, and S.W. is the incumbent of the Dr. Trude Burchardt Professorial Chair of Structural Biology. Supported by a grant from the Philip Klutznick Research Fund.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1161/DC1
Materials and Methods

2 July 2004; accepted 6 October 2004

Earth Tides Can Trigger Shallow Thrust Fault Earthquakes

Elizabeth S. Cochran,[1]* John E. Vidale,[1] Sachiko Tanaka[2]†

We show a correlation between the occurrence of shallow thrust earthquakes and the occurrence of the strongest tides. The rate of earthquakes varies from the background rate by a factor of 3 with the tidal stress. The highest correlation is found when we assume a coefficient of friction of $\mu = 0.4$ for the crust, although we see good correlation for μ between 0.2 and 0.6. Our results quantify the effect of applied stress on earthquake triggering, a key factor in understanding earthquake nucleation and cascades whereby one earthquake triggers others.

For more than a century, researchers have sought to detect the effect on the timing of earthquakes of the gravitational perturbations on Earth from the Moon and Sun (*1*). However, the tidal stresses in most locations are small, and usually it is difficult to ascertain the orientation of the fault plane, which is critical when calculating the effect of the stress variations. Earthquake-tide correlations have been observed to be small or nonexistent in normal crust (*2*–*4*); however, correlations have been shown in shallow, possibly hydrothermal or magma-related areas (*5*, *6*). Here, we take advantage of accurate accounting of ocean tides (*7*) and a large data set of earthquake focal mechanisms with fairly well known fault planes (*8*) to look for a correlation.

We used global earthquakes in the Harvard Centroid Moment Tensor (CMT) catalog (*9*). For each event, we calculated a tidal-stress time series that includes the solid Earth tide and an ocean-loading component (*7*, *10*, *11*). Solid-Earth tides induce stresses only up to 5×10^3 Pa (0.05 bar), whereas in ocean basins, water loading builds stresses up to nearly 5×10^4 Pa (0.5 bar). Both components must be accurately determined to fully resolve tidal influences on the initiation of earthquakes globally. We resolved tidal stresses into normal and shear stress acting on each of the two possible fault planes of the CMT earthquake focal mechanism. Shear failure under compressive stress can be described by the Coulomb criterion, in which a fault fails under a combination of shear and normal stress: $\tau_c = \tau + \mu\sigma_n$, where τ and σ_n are the shear and normal stresses, respectively, and μ is the coefficient of friction. In addition to examining shear and normal stress independently, we tested different values of μ (0.2, 0.4, or 0.6). For each event, we calculated the tidal phase angle (θ) between –180° and 180° (*11*); 0° phase is defined to be at the time of maximum stress that can promote failure, which is extensional for normal stress and in the direction of slip for shear stress (Fig. 1). In addition, we defined the average of the tidal stress amplitudes at the peaks just before and after each earthquake (τ_{pb} and τ_{pa}, respectively) to be the peak tidal stress τ_p.

We focused on a subset of shallow thrust earthquakes with depths of 0 to 40 km because these earthquakes are in regions with the

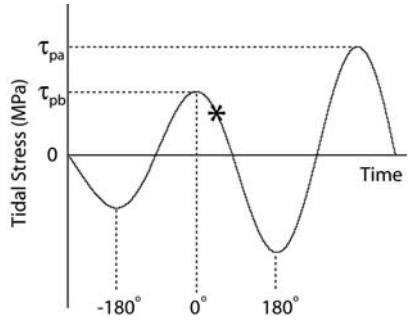


Fig. 1. Schematic diagram of the tidal stress time series spanning 1 day for a hypothetical earthquake (asterisk). Tidal phase is marked, with the maximum Coulomb stress promoting failure defined at 0° phase. The earthquake occurs at $\theta = 45°$. Peak stress amplitudes before and after an event (τ_{pb} and τ_{pa}, respectively) are averaged to determine τ_p.

Table 1. Comparison of coefficients of friction. Data are shown for the 250 events with the highest calculated tidal stress (τ_p) given different values of the coefficient of friction (μ). Binomial is approximated by a Gaussian distribution; *P* values are determined using Schuster's statistical test of data distribution (values below 5% are often considered significantly nonrandom). θ_{peak} is the phase of the peak of a sinusoidal fit to the data. See text for definition of N_{ex}.

Events	Binomial (%)	*P* value (%)	N_{ex} (%)	θ_{peak} (degrees)
$\mu = 0$ (shear)	4.38	10.36	5.6	–22.2
$\mu = 0.2$	0.1439	0.6253	9.6	–1.2
$\mu = 0.4$	0.0032	0.0157	12.8	0.2
$\mu = 0.6$	0.0942	0.3265	10.0	6.0
$\mu = \infty$ (normal)	0.4688	3.677	8.4	15.8

[1]Department of Earth and Space Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095, USA. [2]Department of Geophysics, Graduate School of Science, Tohoku University, Sendai, Miyagi 980-8578, Japan.

*To whom correspondence should be addressed. E-mail: cochran@moho.ess.ucla.edu
†Present address: National Research Institute for Earth Science and Disaster Prevention (NIED), Tsukuba-shi, Ibaraki-ken 305, Japan.

largest tidal stresses; the true fault plane is better known for these events; and shallow events are under lower confining pressures, so tidal stress may be proportionally more influential. In addition, thrust and normal faults have relatively larger tidal stress amplitudes than strike-slip faults because of a larger influence of the ocean loading component. As the ocean moves back and forth in an ocean basin, the additional weight of the water acts to clamp and unclamp dipping faults. In subduction zones, ocean loading tends to be largest and low-angle thrust events are most common. The focal mechanisms for shallow thrust faults show a bimodal distribution of fault dips for fault planes 1 and 2 (Fig. 2A). The shallower-dipping fault planes (plane 1) for the 19 events with τ_p greater than 0.02 MPa (0.2 bar) are aligned with the local geometry of the subduction zone (Fig. 2B), so for this study we assume that to be the true fault plane.

Correlation with the tides is found for shallow-dipping thrust events, assuming μ = 0.2, 0.4, or 0.6 (Fig. 3). Schuster's test, used to find statistical significance of periodicity, indicates less than a 1% probability that the distribution is random for τ_p above about 0.01 MPa (0.1 bar). More events occur during the period of encouraging stress ($-90° < \theta < 90°$) than occur during times of discouraging stress ($-180° < \theta < -90°$ or $90° < \theta < 180°$); the sinusoidal fit to the nonrandom distribution of events peaks at a tidal phase near $\theta = 0°$, which is the phase of tidal stress expected to promote failure (Fig. 4A).

We also performed a binomial test of significance because Schuster's tests may overestimate significance. We estimated the probability of seeing at random the larger fraction of earthquakes we observed to occur during the half of the tidal phase period with encouraging stress, $-90° < \theta < 90°$. If we assume that events occur at random, the number of events that occur with a tidal phase between –90° and 90° should be equal to the number that occur in the other half of the tidal phase range. This simple binomial test gives a probability of >99.99% that earthquakes correlate with tides when tidal stress amplitudes are high (>0.01 MPa) (Fig. 3D). The binomial statistic is calculated assuming that just one choice of parameters was tested; the true odds of finding the correlation at random are expected to be somewhat lower than those obtained here. Tidal correlation with earthquake timing at the highest stress levels is apparent for most choices of coefficient of friction, but the most significant correlation is for μ = 0.4 (Table 1 and Fig. 3D). Assuming μ = 0.4, a minimum binomial probability of 0.0027% of fortuitous correlation (with a corresponding P value of 7.6×10^{-5} is observed for the 255 events that occur at $\tau_p > 0.003$ MPa (*11*).

We see fluctuations in the binomial statistics (Fig. 3D). The most highly correlated 45 events are the most highly stressed, and most are shallow, with 41 events at or above 15 km depth. The next 100 events are not as highly correlated and tend to be deeper, with only 45% of the earthquakes at

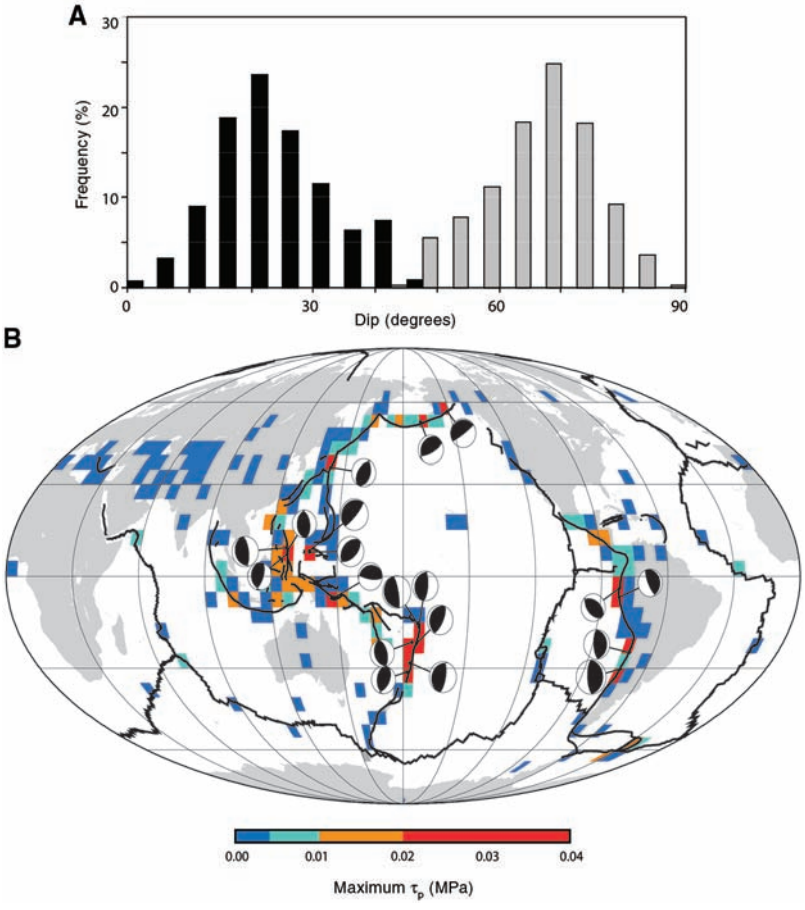


Fig. 2. (**A**) Plot of frequency versus fault dip for 2027 reverse-type earthquakes with hypocenters shallower than 40 km depth. Fault plane 1 (black) and fault plane 2 (gray) dips are from the Harvard CMT catalog focal mechanism solutions. (**B**) Global distribution of maximum τ_p assuming μ = 0.4 in 5° × 5° grids for all $M \geq 5.5$ global shallow thrust earthquakes. Continents are shown in light gray; plate boundaries are shown by solid black lines. Note higher peak tidal stresses (red and orange grids) at continent-ocean margins. Lower-hemisphere focal mechanisms (circular, half-filled symbols) shown for the 19 thrust events with $\tau_p \geq 0.02$ MPa agree with tectonic regime for the shallower-dipping choice of the fault plane.

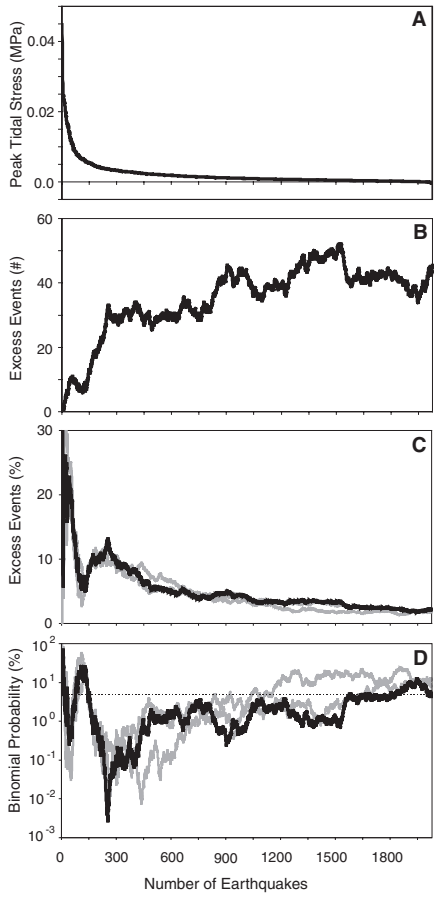


Fig. 3. (**A**) Plot of peak tidal stress τ_p versus total number of events N_{tot}. (**B**) Plot of the number of excess events N_{ex} during times of higher stress versus N_{tot}. (**C**) Plot of percentage of N_{ex} versus N_{tot}. (**D**) Binomial probability versus N_{tot} for μ = 0.2 (black line), 0.4 (gray line), and 0.6 (gray line). Dashed line indicates 5% probability level for reference.

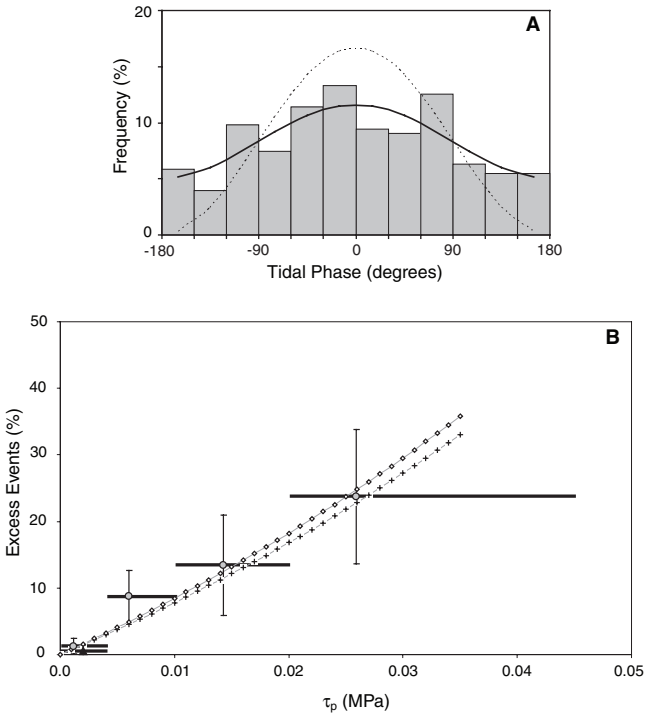


Fig. 4. (**A**) Histogram of frequency of events versus θ for the 255 events with the highest τ_p, assuming $\mu = 0.4$. Solid line is the sinusoidal least-squares fit to the data; dashed line is the Coulomb stress amplitude as a function of tidal phase. (**B**) Percentage of N_{ex} versus τ_p. Values are given in Table 2. Points are located at the mean τ_p; range is indicated by solid horizontal lines. Gray circles are global thrust data; solid triangle at lower left denotes California strike-slip results. Estimated SEs are shown by vertical bars. Crosses and diamonds show the least-squares fit of the data to rate- and state-dependent friction and stress corrosion, respectively.

Table 2. Data summary. Both global thrust events and California strike-slip events (*4*) are shown for various peak Coulomb stress ranges (τ_p) given a coefficient of friction $\mu = 0.4$. *N* is the number of events in each τ_p bin. N_{ex} includes SE.

Data set	τ_p (MPa)	*N*	N_{ex} (%)
Global thrust	>0.02	19	23.7 ± 10.10
Global thrust	0.01 to 0.02	41	13.4 ± 7.52
Global thrust	0.004 to 0.01	155	8.7 ± 3.95
Global thrust	<0.004	1,813	1.2 ± 1.17
California strike-slip	<0.004	27,464	0.60 ± 0.30

or above 15 km depth. We observe a recovery of the tidal correlation as more events are included. This trend suggests a depth dependence to the tidal correlation, but this trend is not statistically significant. Event depth may influence correlation with the tides because confining pressures increase with depth, so there is a comparatively larger tidal influence at shallow depths. However, tidal stress amplitudes decay with depth, so event depth and tidal stress amplitude are not independent.

On the basis of the earthquake-tide correlation observed, we have tried to estimate the tidal stress amplitude required to trigger an earthquake. We define the percent of excess events, N_{ex}, to be the number of additional or excess events in the tidal phase period with encouraging stress, $N_{ex} = [N_{enc} - (N_{tot}/2)]/N_{tot}$, where N_{enc} is the number of events with $-90° < \theta < 90°$ and N_{tot} is the total number of events. For τ_p greater than 0.02 MPa (0.2 bar), N_{ex} is 24%. This corresponds to 74% of the events (14 of 19 earthquakes) occurring in 50% of the phase time of encouraging stress ($-90° < \theta < 90°$). Successively lower peak tidal stresses show lower tidal correlations, indicating less effective earthquake triggering (Table 2 and Fig. 4B). Standard errors, calculated using data variance, are large because of small numbers of earthquakes in each τ_p bin. In addition to the global data set, 27,464 strike-slip events from California were found to have roughly 1 to 2% more events during times of encouraging tidal stress of <0.004 MPa (Fig. 4B).

A previous study using global data hinted at an earthquake-tide correlation, suggesting that reverse and normal earthquakes correlate either with the shear stress or the trace of the stress tensor J1 (*7*). In addition, a few regional studies (*5*, *12*–*14*) in areas with a large ocean-loading component of tidal stress have observed triggering of earthquakes or volcanic tremor. Most often, triggered events have been shallow, normal-type faulting along axial ridges, such as Juan de Fuca (*5*, *12*), perhaps associated with hydrothermal circulation. A recent study of seismicity in Japan showed earthquake-tide correlation in subregions, with the best correlations found in regions that experienced a large earthquake (*15*). Our study shows a statistically significant increase in triggering of globally distributed large tectonic events ($M > 5.5$) with increasing tidal stress amplitudes (Fig. 4).

Tidal stress amplitudes required to trigger earthquakes are similar to thresholds suggested for static and dynamic triggering of aftershocks of 0.003 MPa (*16*–*18*). A study of triggering after the *M*7.4 Landers event, for example, showed that aftershocks were triggered by stress increases greater than 0.01 MPa (*16*). In addition, the observed trend of increased triggering with higher imposed tidal stress can be well fit to friction theories of rate- and state-dependent friction and stress corrosion (*11*, *19*, *20*).

References and Notes

1. D. Emter, in *Tidal Phenomena*, H. Wilhelm, W. Zurn, H.-G. Wenzel, Eds., vol. 66 of *Lecture Notes in Earth Sciences* (Springer-Verlag, Berlin, 1997), pp. 293–310.
2. S. Hartzell, T. Heaton, *Bull. Seismol. Soc. Am.* **79**, 1282 (1989).
3. J. E. Vidale, D. C. Agnew, M. J. S. Johnston, D. H. Oppenheimer, *J. Geophys. Res.* **103**, 24567 (1998).
4. J. E. Vidale, D. Agnew, D. Oppenheimer, C. Rodriquez, H. Houston, *Eos* **79** (suppl.), F641 (1998).
5. M. Tolstoy, F. L. Vernon, J. A. Orcutt, F. K. Wyatt, *Geology* **30**, 503 (2002).
6. M. O. Saar, M. Manga, *Earth Planet. Sci. Lett.* **214**, 605 (2003).
7. S. Tanaka, M. Ohtake, H. Sato, *J. Geophys. Res.* **107**, 2211 (2002).
8. A. M. Dziewonski, T.-A. Chou, J. H. Woodhouse, *J. Geophys. Res.* **86**, 2825 (1981).
9. The entire data set consists of the 9350 global earthquakes of *M*5.5 or greater from 1977 to 2000 in the Harvard CMT catalog. These events include 2823 reverse, 1040 normal, 3597 strike-slip, and 1890 oblique-type faulting events (*7*).
10. K. Matsumoto, T. Takanezawa, M. Ooe, *J. Oceanogr.* **56**, 567 (2000).
11. See supporting data on *Science* Online.
12. W. S. D. Wilcock, *Geophys. Res. Lett.* **28**, 3999 (2001).
13. S. Tanaka, M. Ohtake, H. Sato, *Geophys. Res. Lett.* **29**, 10.1029/2002GL015386 (2002).
14. S. I. S. Custodio, J. F. B. D. Fonseca, N. F. d'Oreye, B. V. E. Faria, Z. Bandomo, *Geophys. Res. Lett.* **30**, 1816 (2003).
15. S. Tanaka, M. Ohtake, H. Sato, *Earth Planets Space* **56**, 511 (2004).
16. R. S. Stein, *Nature* **402**, 605 (1999).
17. J. L. Hardebeck, J. J. Nazareth, E. Hauksson, *J. Geophys. Res.* **103**, 24427 (1998).
18. R. S. Stein, *Sci. Am.* **288**, 72 (January 2003).
19. J. H. Dieterich, *Tectonophysics* **144**, 127 (1987).
20. I. Main, *Geophys. J. Int.* **139**, F1 (1999).

21. We thank P. Bird, E. Brodsky, J. Dieterich, H. Houston, D. Jackson, C. Scholz, R. Stein, and three anonymous reviewers for helpful comments and suggestions and K. Matsumoto for making available the ocean tide model NAO.99b. Supported by NSF grant 0125732, an NSF graduate fellowship (E.S.C.), and grant-in-aid for JSPS fellows 07136 (S.T.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/1103961/DC1
Materials and Methods
References

11 August 2004; accepted 7 October 2004
Published online 21 October 2004;
10.1126/science.1103961
Include this information when citing this paper.

Geochemical Precursors to Volcanic Activity at Mount St. Helens, USA

Kim Berlo,[1]* Jon Blundy,[1] Simon Turner,[1,2] Kathy Cashman,[3] Chris Hawkesworth,[1] Stuart Black[4]

The importance of the interplay between degassing and crystallization before and after the eruption of Mount St. Helens (Washington, USA) in 1980 is well established. Here, we show that degassing occurred over a period of decades to days before eruptions and that the manner of degassing, as deduced from geochemical signatures within the magma, was characteristic of the eruptive style. Trace element (lithium) and short-lived radioactive isotope (lead-210 and radium-226) data show that ascending magma stalled within the conduit, leading to the accumulation of volatiles and the formation of lead-210 excesses, which signals the presence of degassing magma at depth.

On 18 May 1980, an earthquake-triggered landslide exposed a subsurface lava "cryptodome" at Mount St. Helens that led to its eruption, first in the form of a violent lateral blast and subsequently as a 20-km-high Plinian column and associated pyroclastic flows (*1*). The eruption followed 2 months of precursory activity in the form of seismic events, phreatomagmatic explosions, and deformation, which are attributed to the emplacement of the cryptodome. Cryptodome samples are typically dense, with a highly crystalline groundmass of rapidly grown plagioclase microlites, crystals that provide evidence of substantial crystallization produced by decompression and degassing (*2*, *3*). The 18 May Plinian eruption was followed by a series of smaller explosive eruptions during the summer of 1980 and by 6 years of intermittent effusive activity, a pattern interpreted to reflect gradually decreasing rates of magma supply (*4*, *5*). A pronounced increase in groundmass crystallinity of the erupted products during this time period testifies to the increasing importance of degassing-induced crystallization (*3*, *4*, *6*). Where, when, and how rapidly this crystallization occurred, and the physical effects of degassing on the evolution of the eruption, are more poorly constrained.

The short-lived isotopes ^{226}Ra ($t_{1/2}$ = 1599 years) and ^{210}Pb ($t_{1/2}$ = 22.6 years) can be used to constrain the time scale of degassing, because the key intermediate isotope ^{222}Rn ($t_{1/2}$ = 3.8 days) is volatile ($^{226}\mathrm{Ra} \rightarrow {}^{222}\mathrm{Rn} \xrightarrow{{}^{218}\mathrm{Po}\rightarrow{}^{214}\mathrm{Pb}\rightarrow{}^{214}\mathrm{Bi}\rightarrow{}^{214}\mathrm{Po}} {}^{210}\mathrm{Pb}$). During degassing of major volatile species such as H_2O, SO_2, and CO_2, ^{222}Rn will diffuse into the gas bubbles, thereby removing ^{222}Rn and leading over time to a ^{210}Pb deficit in the magma from which volatiles escape. Whole-rock samples from the 18 May Plinian phase have a minimum (^{210}Pb/^{226}Ra) of 0.84 (parentheses denote activity), which indicates a minimum of 5 years of continuous pre-eruptive degassing (*7*). However, samples from the cryptodome and several pumices erupted in the summer of 1980 have ^{210}Pb excesses, that is, (^{210}Pb/^{226}Ra) > 1 (Fig. 1A), which have been reported in very few studies (*7*, *8*). Contrary to previous suggestions based on an earlier study (*8*), we suggest that there is a trend of increasing (^{210}Pb/^{226}Ra) with time during the summer of 1980 (Fig. 1A). In contrast to the formation of ^{210}Pb deficits where exsolved gases leave the system, resulting in the loss of ^{222}Rn, ^{210}Pb excesses require gas accumulation in the conduit, allowing for the decay of ^{222}Rn to ^{210}Pb while in contact with the melt, but remote from the site of degassing. Thus, ^{210}Pb excesses suggest the presence of a deeper degassing reservoir supplying volatiles to shallow stalled magma. While the creation of significant ^{210}Pb deficits requires years (*7*), ^{210}Pb excesses may be established on much shorter time scales. The magnitude of these excesses depends on the flux of radon relative to the rate of ^{222}Rn decay. A simple calculation assuming secular equilibrium between (^{226}Ra) and (^{222}Rn) in the deeper magma reservoir, complete degassing of radon, and sufficient time (~1 month) for all ^{222}Rn in the gas to decay to ^{210}Pb shows that the volume of degassing magma needs to be roughly 10 times that of the shallow reservoir to produce a (^{210}Pb/^{226}Ra) of 1.4. If there is less time available for ^{222}Rn to decay, the degassed volume needs to be larger.

Additional constraints on the time scales of degassing can be obtained from plagioclase phenocrysts, which record changes in the magmatic environment by changing composition. Although the plagioclase phenocrysts in Mount St. Helens dacite are variable in texture and composition within any one sample, all samples display the same textures and compositional range. This is reflected in the similar Ba, Sr, Ti, and Mg contents of plagioclase phenocrysts from different eruptions. The striking exception is Li (Fig. 2). The Li concentration of plagioclase phenocrysts in 18 May pumices is generally low, whereas those erupted before and after 18 May have higher Li concentrations (Fig. 1B and Fig. 2). Li diffusion in melt (*9*) and plagioclase (*10*) is extremely rapid. As Li partitions into vapor relative to melt at depth (*11*), low Li concentrations in 18 May plagioclase are consistent with degassing, as inferred from ^{210}Pb deficits in the associated whole rocks.

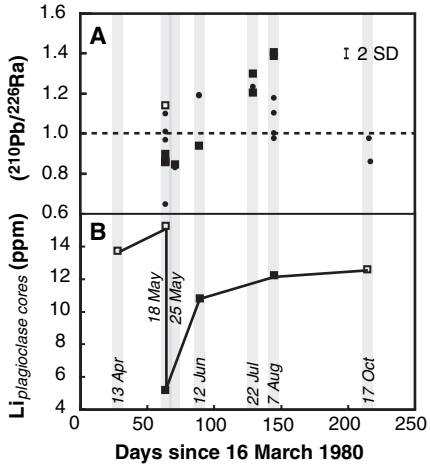


Fig. 1. Whole-rock variation of (^{210}Pb/^{226}Ra) and Li concentration in plagioclase during summer 1980. (**A**) The earliest sign of a pending eruption was recorded on 16 March (*1*); hence, this is taken as t = 0. Secular equilibrium is indicated by the dashed line (^{210}Pb/^{226}Ra) at unity. Black squares indicate pumice samples, open squares are dome samples, and dots are data from Bennett *et al*. (*8*). (^{210}Pb/^{226}Ra) is measured by a combination of gamma and alpha spectrometry (*23*). The vertical lines with dates indicate eruptions. A maximum error bar (2 SD) is in the top right corner. (**B**) The lithium concentration of plagioclase phenocrysts increases with time and follows a pattern similar to (^{210}Pb/^{226}Ra). This figure shows the variation in Li concentration in the cores of phenocrysts [defined here as having Ba concentrations between 25 and 80 parts per million (ppm); see Fig. 2]. After detailed scanning electron microscopy study, two to five phenocrysts of each eruption were selected to be analyzed by secondary ion mass spectrometry; see (*23*) for details. Plagioclase compositions for all samples lie within the range An_{38} to An_{82}.

[1]Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Rd, Bristol BS8 1RJ, United Kingdom. [2]GEMOC, Department of Earth and Planetary Sciences, Macquarie University, Sydney, NSW 2109, Australia. [3]Department of Geological Sciences, 1272 University of Oregon, Eugene, OR 97403–1272, USA. [4]Department of Archaeology, School of Human and Environmental Sciences, University of Reading, Whiteknights, PO Box 227, Reading RG6 6AB, UK.

*To whom correspondence should be addressed. E-mail: Kim.Berlo@bristol.ac.uk

Melt inclusions in phenocrysts can be used to study variation in the concentrations of volatile elements during magma ascent. H_2O concentration depends on the partial pressure of H_2O, pH_2O (*12*), and thus, in volatile-saturated CO_2-poor magma, it can be used as a monitor of the pressure at which the melt inclusion was last in contact with the magma. The variable H_2O concentrations in melt inclusions from Mount St. Helens indicate that they have been sealed off from the magma at different depths (Fig. 3). The compositional variation of melt inclusions details the evolution of the melt while undergoing degassing and crystallization. The melt inclusions are crystal free, and their MgO contents indicate minimal post-entrapment crystallization (*13*). The variation of Li with H_2O (Fig. 3) shows that at $pH_2O \approx$ 125 MPa, the melt inclusions record Li enrichment. For all explosive eruptions after 18 May, Li enrichment is confined to this depth (Fig. 3, scenario 3), whereas for effusive eruptions the magnitude of Li enrichment gradually decreases during ascent (Fig. 3, scenario 1).

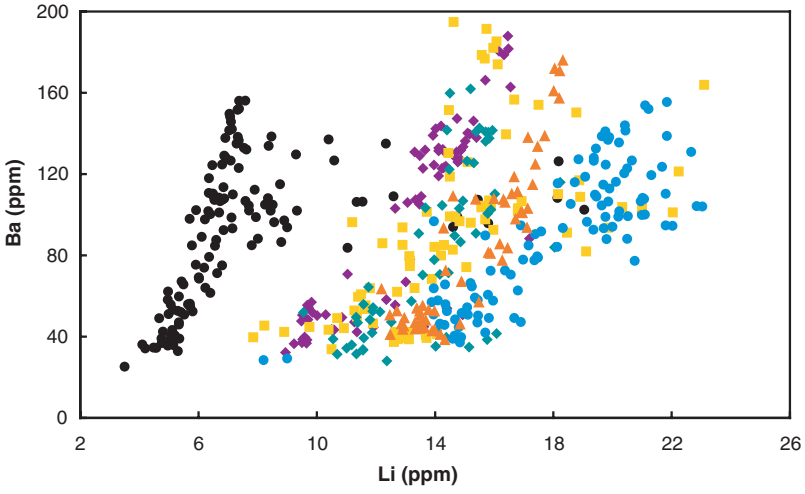


Fig. 2. Ba and Li concentrations of plagioclase phenocrysts. Ba and Li variation in plagioclase phenocrysts erupted on 13 April (orange triangles), 18 May cryptodome (blue circles), 18 May climactic eruption (black circles), 12 June (purple diamonds), 7 August (yellow squares), and 17 October dome (green diamonds). For details on the analyses and for data, see (*23*).

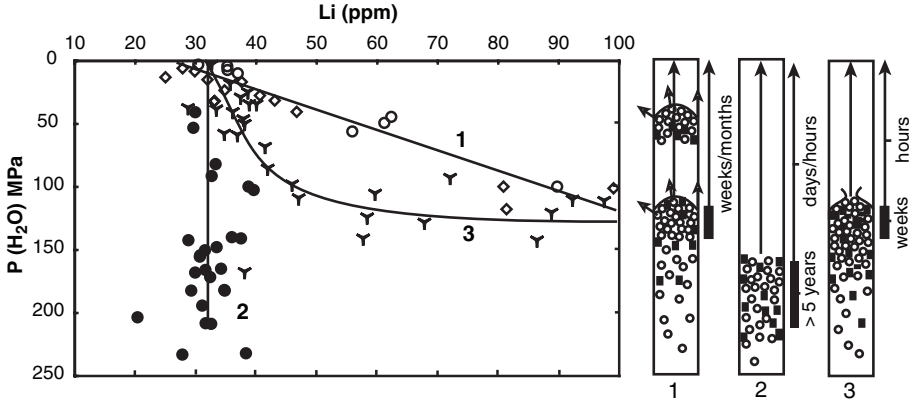


Fig. 3. H_2O and Li concentrations of melt inclusions in phenocrysts from various eruptions during the summer of 1980. H_2O concentration is taken as an indication of the minimum pressure at which the melt inclusions last equilibrated with the melt. H_2O saturation pressures were calculated using VolatileCalc (*12*). Melt inclusions from the 18 May pumice were mainly trapped at high pressures and low Li concentrations (circles, scenario 2). Melt inclusions from explosive eruptions after 18 May record high Li concentrations at intermediate depths (tripods, scenario 1), followed by a rapid decrease in Li concentration during ascent. The cryptodome and October dome (open circles and diamonds, respectively) have melt inclusions recording the same Li enrichment event as seen in the pumices, but during their ascent the Li concentration in the melt was buffered by a continuous gas flow. Li enrichment is inferred to result from gas fluxing, the same process responsible for the formation of ^{210}Pb excesses (Fig. 1). Deep degassing magma supplies gas with Rn and Li to shallow conduit magma. If magma stalls, gases can accumulate in the conduit. At this point, the impermeable cap is breached either suddenly (scenario 3), followed by rapid magma ascent, or more slowly (scenario 1) by leaking gas while ascending slowly as a semipermeable plug. Magma erupted during the 18 May Plinian phase did not stall, but ascended rapidly.

The combination of Li enrichment in plagioclase and melt inclusions for all erupted rocks except 18 May pumices suggests that Li is added at 4- to 5-km depth beneath the volcano. The combined data set indicates that Li and ^{210}Pb (or ^{222}Rn) are transferred from a deeper reservoir (as erupted during the Plinian phase of 18 May) to shallow-level magma erupted subsequently. As Rn and Li are expected to partition into the vapor relative to the melt at high pressure, transfer by volatiles is the most likely mechanism for both ^{210}Pb excesses and Li enrichment. Magma stalling at shallow levels where gases from the deeper magma are accumulated can thus acquire an elevated Li concentration and ^{210}Pb excess. Time scales of ^{210}Pb and Li enrichment differ. Because of its extremely high diffusivity in melt and plagioclase (*9*, *10*), Li enrichment depends mainly on the gas flux. Conversely, ^{210}Pb enrichment also relies on the decay of ^{222}Rn. Thus, eruptions with high Li concentrations and ^{210}Pb deficits (e.g., June eruption, Fig. 1) imply that degassed magma has been stored shallowly, but not long enough to develop ^{210}Pb excess. An additional constraint for formation of ^{210}Pb excesses is presented by the seismic and SO_2 records, which suggest magma ascent 9 days prior to the 12 June eruption (*1*, *14*), suggesting that 9 days is insufficient to create ^{210}Pb excesses. Figure 3 shows that during subsequent ascent and eruption of shallow magma, Li in the melt is lost to the volatile phase (post–18 May explosive eruptions) or buffered by the gas flow (*15*) (effusive eruptions). In contrast, the Plinian 18 May pumices show different behavior with respect to Li and ^{210}Pb: Their low-Li plagioclase cores and ^{210}Pb deficits indicate that the magma has lost volatiles earlier and did not stall during ascent.

The ^{210}Pb and Li data require gas accumulation and transfer of volatiles from vapor to melt at low pressure. Vapor transport in magma may occur by bubble ascent, convection (*16*), and/or shearing and fracturing (*15*). Thus, volcanic gases can accumulate in permeable foam (*17*–*19*) underneath an impermeable cap or simply in a region where magma stalls because of conduit geometry or favorable crystallization conditions. The cause of transfer of Li and ^{210}Pb from the vapor to the melt is unclear, but, because both form chloride complexes, it may be related to the exsolution of an immiscible brine phase at shallow pressure (*20*), as widely proposed for ore-forming hydrothermal systems (*21*). Whatever the transfer mechanism, the only records of these changes to survive are the melt inclusions that are sealed off progressively during ascent, the Li concentration in plagioclase (if ascent is fast enough to prevent

diffusive reequilibration), and the disequilibrium between ^{226}Ra and ^{210}Pb. Once ^{210}Pb excess is formed, reestablishing secular equilibrium between ^{226}Ra and ^{210}Pb would require about 100 years; thus, the ^{210}Pb excesses formed by accumulation of gases will persist. Plagioclase grows during decompression (under H_2O-saturated conditions), thus sealing off the melt inclusions shortly after gas is lost. The available experimental data show that Li diffuses more rapidly than all other trace elements in melt (*9*) and plagioclase (*10*), which suggests that Li can be diffusively homogenized in both phases on time scales of hours.

In the context of Mount St. Helens, our results provide a detailed picture of magma and gas movement during 1980. Specifically, the cryptodome followed a two-stage decompression path interrupted at a depth of 4- to 5-km where magma stalled and gas accumulated before further ascent. During magma ascent, continued gas fluxing through a semipermeable magmatic foam buffered the Li concentration of the melt and supplied Rn (Fig. 3). Evidence for breaching of the impermeable cap and gas flow comes from phreatomagmatic eruptions and steam venting in March and April (*2*). Magma erupted during the Plinian phase of the eruption ascended rapidly from a depth of more than about 7 km without stalling (*22*). Before the post–18 May explosive eruptions, magmas stalled at 4- to 5-km before the impermeable cap was breached and magma ascended rapidly. A dome sample from October 1980 also has high Li melt inclusions at $pH_2O < 125$ MPa, which suggests that magma ascended more slowly than during the preceding explosive eruptions and, as with the cryptodome, Li concentrations in the melt were buffered. The increase of (^{210}Pb/^{226}Ra) during the summer of 1980 correlates with a decrease in magma ascent rate (*4*, *5*). This suggests that the trend of increasing (^{210}Pb/^{226}Ra) with time reflects prolonged stalling of magma at 4- to 5-km depth as the eruption intensity waned.

Only recently has attention been turned to the complexity of shallow conduit processes and their link to eruption style. Our data shed light on processes occurring on time scales of years to hours before an eruption and, as such, may provide an aid in interpreting observations from established monitoring techniques, for example, gas emissions and seismic surveys. ^{210}Pb excesses are coupled to ^{210}Pb deficits deeper in the system. Thus, magma erupting with ^{210}Pb excess requires the presence of degassing magma at depth.

References and Notes

1. P. W. Lipman, D. R. Mullineaux, Eds., *U.S. Geol. Surv. Prof. Paper* **1250**, 93 (1981).
2. K. V. Cashman, R. P. Hoblitt, *Geology* **32**, 141 (2004).
3. K. Cashman, *Contrib. Mineral. Petrol.* **109**, 431 (1992).
4. C.-H. Geschwind, M. J. Rutherford, *Bull. Volcanol.* **57**, 356 (1995).
5. R. Scandone, S. D. Malone, *J. Volcanol. Geotherm. Res.* **23**, 239 (1985).
6. J. Blundy, K. Cashman, *Contrib. Mineral. Petrol.* **140**, 631 (2001).
7. P.-J. Gauthier, M. Condomines, *Earth Planet. Sci. Lett.* **172**, 111 (1999).
8. J. T. Bennett, S. Krishnaswami, K. K. Turekian, W. G. Melson, C. A. Hopson, *Earth Planet. Sci. Lett.* **60**, 61 (1982).
9. F. M. Richter, A. M. Davis, D. J. DePaolo, E. B. Watson, *Geochim. Cosmochim. Acta* **67**, 3905 (2003).
10. B. J. Giletti, T. M. Shanahan, *Chem. Geol.* **139**, 3 (1997).
11. J. D. Webster, J. R. Holloway, R. L. Hervig, *Econ. Geol.* **84**, 116 (1989).
12. S. Newman, J. B. Lowenstern, *Comput. Geosci.* **28**, 597 (2002).
13. J. Blundy, K. Cashman, in preparation.
14. T. Casadevall *et al.*, *Science* **221**, 1383 (1983).
15. A. C. Rust, K. V. Cashman, P. J. Wallace, *Geology* **32**, 349 (2004).
16. D. S. Stevenson, S. Blake, *Bull. Volcanol.* **60**, 307 (1998).
17. B. E. Taylor, J. C. Eichelberger, H. R. Westrich, *Nature* **306**, 541 (1983).
18. J. C. Eichelberger, C. R. Carrigan, H. R. Westrich, R. H. Price, *Nature* **323**, 598 (1986).
19. C. Klug, K. C. Cashman, *Bull. Volcanol.* **58**, 87 (1996).
20. S. Signorelli, M. R. Carroll, *Geochim. Cosmochim. Acta* **64**, 2851 (2000).
21. H. L. Barnes, Ed., *Geochemistry of Hydrothermal Ore Deposits* (Wiley, New York, ed. 3, 1997).
22. M. J. Rutherford, P. M. Hill, *J. Geophys. Res.* **98**, 19,667 (1993).
23. Methods and data tables are available as supporting material on *Science* Online.

24. During this research, K.B. was supported by a University of Bristol scholarship; J.B. and S.T. acknowledge support from the Royal Society. W. Melson at the Smithsonian Institution provided some of the samples used for this study. Other samples were collected with the generous support of J. Pallister and M. Clynne during field work. We also thank S. Kasemann (Natural Environment Research Council Ion Microprobe Facility, Edinburgh) for assistance with secondary ion mass spectrometry analysis.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1103869/DC1
Materials and Methods
Tables S1 to S3

10 August 2004; accepted 1 October 2004
Published online 14 October 2004;
10.1126/science.1103869
Include this information when citing this paper.

Ventilation of the Glacial Deep Pacific Ocean

Wallace Broecker,[1]* Stephen Barker,[1] Elizabeth Clark,[1] Irka Hajdas,[2] Georges Bonani,[2] Lowell Stott[3]

Measurements of the age difference between coexisting benthic and planktic foraminifera from western equatorial Pacific deep-sea cores suggest that during peak glacial time the radiocarbon age of water at 2-kilometers depth was no greater than that of today. These results make unlikely suggestions that a slowdown in deep-ocean ventilation was responsible for a sizable fraction of the increase of the ratio of carbon-14 (^{14}C) to carbon in the atmosphere and surface ocean during glacial time. Comparison of ^{14}C ages for coexisting wood and planktic foraminifera from the same site suggests that the atmosphere to surface ocean ^{14}C to C ratio difference was not substantially different from today's.

Hughen *et al.* (*1*) present a strong case that during the last glacial maximum (LGM) (i.e., 22,000 to 16,000 calendar years ago) the ^{14}C to C ratio in the atmosphere and surface ocean was 375 ± 25‰ higher than that for preindustrial time. The obvious explanation for this increase is that Earth's magnetic field was on the average weaker, allowing more cosmic rays to reach our atmosphere. Hughen *et al.* (*1*) contend, however, that although the field was weaker, the consequent increase in ^{14}C production was insufficient to explain the entire observed ^{14}C increase. To explain the remainder, these authors call on a sizable reduction in deep-sea ventilation rate and a possible reduction in shallow marine carbonate deposition (i.e., reef growth). Muscheler *et al.* (*2*) reach a similar conclusion by another route. They make use of the ^{10}Be measurements on the Summit Greenland ice core to reconstruct past production rates of cosmogenic isotopes and conclude that the excess ^{14}C present in the atmosphere and upper ocean cannot be accounted for by higher production alone. They also call on reduced deep-ocean ventilation in an attempt to account for at least part of the remaining radiocarbon increase.

Were the deep sea to have been ventilated at a slower rate during glacial time, then the ^{14}C to C ratio difference between upper-ocean and deep-ocean carbon must have been larger than today's. As shown in Table 1, raising the atmosphere and surface ocean ^{14}C to C ratio by 200 per mil requires a large increase in the difference between the ^{14}C to

[1]Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W/Post Office Box 1000, Palisades, NY 10964, USA. [2]AMS 14C lab, ETH Hoenggerberg HPK H27 and H30, CH-8093 Zurich, Switzerland. [3]University of Southern California, 3651 Trousdale Parkway, Los Angeles, CA 90089, USA.

*To whom correspondence should be addressed. E-mail: broecker@ldeo.columbia.edu

C ratio for surface and deep-sea inorganic carbon. As this difference is potentially recorded by the ^{14}C age difference between coexisting benthic and planktic foraminifera shells in glacial-age sediments, it is possible to check the prediction by Hughen *et al*. (*1*) and Muscheler *et al*. (*2*) that deep-sea ventilation rates were slower during glacial time.

In a recently published paper (*3*), a rather bleak summary was presented of published attempts to establish the age difference between coexisting benthic and planktic foraminifera in glacial-age sediments from the deep Pacific Ocean. The published results for age difference based on individual benthic-planktic pairs range from as low as a few hundred to as high as 3500 years. New results from the eastern equatorial Pacific presented by Broecker *et al*. (*3*) also yield a wide range of ages for coexisting planktic foraminifera. Further, as acid leaching decreased the age of the residual calcite in these samples, the suggestion is that secondary contaminant calcite must be present. Taken together, the average age difference as a function of water depth for the entire ensemble of benthic-planktic pairs is consistent with that for preindustrial time. However, this average has meaning only if it can be assumed that the biases responsible for the very large range in results are random (i.e., on the average not different for benthics than for planktics). The important point is that only a few of these published age differences are large enough to be consistent with the requirement that the extra atmosphere and upper ocean ^{14}C was the result of a slowdown in deep-sea ventilation.

Having found that foraminifera from glacial-age sediments in the eastern equatorial Pacific appear to be contaminated with secondary calcite, we turned to sediments from the western equatorial Pacific in hopes of avoiding this problem. We obtained 50-g samples from core MD98-2181 from the Morotai Basin at 6°N and 126°E, a location just to the south of the island of Mindanao at a depth of 2.1 km, and from core MD97-2138 from the open Pacific at 1°S and 146°E just to the north of the Admiralty Islands at a depth of 1.9 km (*4*). The planktics in these samples yielded ^{14}C ages that respectively correspond to calendar ages of about 19,000 and 21,500 years. The results are listed in Table 2. Because of the problems previously encountered for the eastern equatorial Pacific samples, the first set of runs was limited to planktics. Excellent agreement was achieved in both cores for four planktic species. Based on this success, we conducted a second run in which one of the planktic species was repeated along with the mixed benthics. Subsequent measurements on two additional samples, one from above and one from below the original one, yielded consistent results.

As a bonus, in the first sample from MD98-2181, we were surprised to find a number of small pieces of wood (*5*). The two largest pieces were included in the second run. The age difference of 560 ± 150 years between the wood and planktic foraminifera is consistent with expectation. Measurements on corals from the western equatorial Pacific indicate that before the onset of nuclear testing the $\Delta^{14}C$ for surface waters in that region was –58 ± 2‰ (*6*, *7*), which corresponds to a reservoir age of 480 years (*8*). As pointed out by Bard (*9*), other things being equal, this difference should be inversely related to the atmosphere's CO_2 content. Taking the LGM value to be 195 μatm and the preindustrial to be 280, the prediction would be that during the LGM the age difference would have been 630 years.

Water-column profiles of $\Delta^{14}C$ in the western equatorial Pacific yield a value of about –220‰ at 2000-m water depth (*10*, *11*). Using the coral-based value of –58‰ for preindustrial surface water in that region, this corresponds to an age of 1520 years. In today's ocean below 1500 m, there is little change in $\Delta^{14}C$ with water depth.

The benthic-planktic ^{14}C age differences obtained for the Morotai core are 1325, 1170, and 1325 years (Figs. 1 and 2). For the Admiralty core, the differences are, respectively, 1954 and 1640 years. Because atmospheric

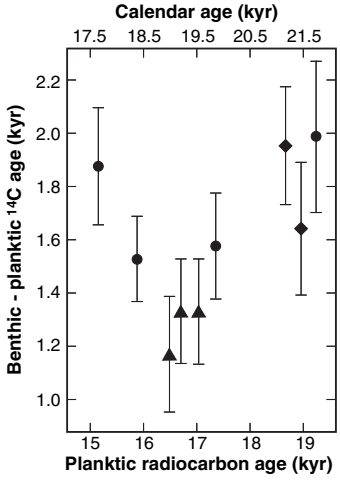


Fig. 1. Benthic-planktic age differences as a function of planktic age for three sites in the western tropical Pacific. The circles represent the results for the South China Sea (*15*), the triangles those for the Morotai Basin, and the diamonds those for Admiralty Island. The upper scale gives the approximate calendar age for the planktics.

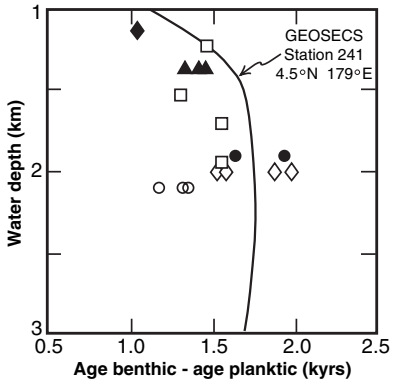


Fig. 2. Summary of LGM ventilation times based on benthic-planktic age differences for cores from the western Pacific. The open squares are for cores from the Sea of Okhotsk (*16*), the solid triangles are for a core from off Japan (*17*), the open diamonds are from a core at a depth of 2.7 km in the South China Sea (sill depth 2.0 km) (*15*), the solid circles are for the Admiralty Island core, and the open circles are for the Morotai Basin core. The solid diamond is for a ^{230}Th-dated benthic coral from the Drake Passage (*18*). Not shown are results obtained on cores stored dry in the Lamont-Doherty repository (*21*) and those from off New Zealand (*22*). The solid line is based on measurements of water samples collected as part of the Geochemical Ocean Sections Study (GEOSECS) survey.

Table 1. Calculation based on a simple box ocean model of the decrease in the ^{14}C to C ratio (relative to that for the preindustrial atmosphere) for deep inorganic carbon required to raise the upper ocean and atmosphere $\Delta^{14}C$ with no change in ^{14}C production. Also given is the increase in the radiocarbon age difference between coexisting benthic and planktic foraminifera. Case 1: Two-box ocean; distribution of inorganic carbon between atmosphere, surface, and deep reservoirs = 0.75:1:38. Surface ocean reservoir age is held at 500 years. Case 2: Three-box ocean; as before, except the deep ocean is now divided into intermediate and deep boxes. The ^{14}C age difference between the surface and intermediate boxes is held at 1500 years.

Atmospheric ^{14}C:C	Surface ^{14}C:C	Intermediate ^{14}C:C	Deep ^{14}C:C	Total ^{14}C:C	Benthic-planktic ^{14}C age (years)
		Case 1. Two-box ocean; surface:deep = 1:38			
1.00	0.940		0.780	0.788	1500
1.10	1.034		0.775	0.788	2312
1.20	1.128		0.771	0.788	3057
		Case 2. Three-box ocean; surface:intermediate:deep = 1:19:19			
1.00	0.940	0.780	0.780	0.788	1500
1.10	1.034	0.858	0.693	0.788	3214
1.20	1.128	0.936	0.606	0.788	4989

and surface ocean $\Delta^{14}C$ was relatively stable during the period 25,000 to 19,000 calendar years ago, it is not necessary to account for changes in atmospheric $\Delta^{14}C$ when interpreting these benthic-planktic offsets (*12*). Further, as the deep sea is continuously ventilated, the ^{14}C to C ratio for the deep sea reflects an average over a time interval of on the order of a millennium rather than a single point in time. The results suggest that waters at 2000 m during the LGM in the western equatorial Pacific were no older with respect to surface waters than today and could actually have been slightly better ventilated at around 19,000 calendar years ago. As these results represent only the mid-depth of the Pacific Ocean, it is possible that the deep high-salinity layer documented for glacial time by Adkins *et al*. (*13*) had a much lower ^{14}C to C ratio. However, as more than half the ocean lies above 2000 m, if the discrepancy pointed out by Hughen *et al*. (*1*) and Muscheler *et al*. (*2*) is to be explained by a slower deep-sea ventilation rate during glacial time, the age of this deep carbon in deeper waters would have to have been at least 4000 years.

Although no reliable ^{14}C data are available for these deeper waters, ^{13}C measurements of glacial-age benthic foraminifera from equatorial zone cores from 1.6 to 4.5 km depth show no particular trend (*14*). If the deep waters were ventilated much more slowly, one would expect that the benthic ^{13}C would have been more negative than that at 2 km. If anything, the data suggest slightly better ventilated waters below 3 km in the glacial Pacific.

The benthic-planktic age differences presented here are consistent with those published for the South China Sea where concordant ages between multiple planktic species were also obtained (*15*). Taken together, the results from these three locales (representing water depths of 2.0 ± 0.1 km in the western Pacific) suggest a dip in ventilation age centered at 19,000 calendar years (see Fig. 1). However, verification of this dip will require further measurements.

LGM benthic–planktic age differences from the Sea of Okhotsk (*16*) and from off Japan (*17*), as well as the estimate based on a ^{230}Th-dated benthic coral from the Drake Passage (*18*), are consistent with those presented here (see Fig. 2). Taken together they suggest that, if anything, LGM ventilation rates were on the average somewhat less than today's.

Clearly, many more measurements will be needed to create an adequate radiocarbon inventory for the LGM ocean. As it is now clear that the formation of secondary calcite is a problem, it will be essential to obtain cross-checks by analyzing several coexisting planktic species. Further, as bioturbation, coupled with opposing gradients in the abundances of planktic and benthic foraminifera can give rise to sizable biases in the reconstructions of surface to deep radiocarbon differences (*19*, *20*), cores with high accumulation rates must be sought. This latter restriction eliminates most open ocean sites and focuses attention on sites located along continental margins. Another problem is that calcite dissolution limits the water depth at which such measurements can be conducted. In the Pacific and Indian Oceans, few calcitic shells are preserved at depths greater than 4400 m.

If it turns out that the reduced storage of ^{14}C in the deep sea is not the answer, then what are the alternatives? The Muscheler *et al*. (*2*) argument rests on the assumption that the ratio of the accumulation of ^{10}Be to accumulation of snow on the Greenland ice cap was not substantially different during the LGM than during the Holocene. Considering that the accumulation rate of ^{10}Be on Greenland is only one-third of the mean global production rate, this assumption rests on shaky ground. If, for example, the accumulation of ^{10}Be relative to snow was lower during glacial time, the discrepancy between the reconstructed ^{14}C and ^{10}Be production rates would disappear.

Hughen *et al*.'s (*1*) contention that variations in Earth's magnetic field strength are not large enough to explain the high LGM ^{14}C to C ratios for the atmosphere and surface ocean present a more serious challenge. These authors point out, however, that unexplained differences between simulated ^{14}C and their observations may be the result of uncertainties in the relation between ^{14}C production rate and geomagnetic intensity. Increasing the database of reliable paleo deep-ocean ventilation ages will provide important constraints for such reconstructions in the future (*21*, *22*).

Table 2. Radiocarbon ages for two western equatorial Pacific deep-sea cores. The uncertainty quoted for mean planktic ages is the standard deviation about the mean age.

Depth (cm)	Species	^{14}C age (years)	1 sigma error (years)
	Morotai Basin MD98-2181 6°N 126°E 2.1 km		
1262–1268			
	Pulleniatina obliquiloculata	16,650	110
	Neogloboquadrina dutertrei	16,800	95
	Mix benthics	18,050	130
	Mean planktics	16,725	150
	Δ B-P	1,325	200
1270–1276			
	Wood 1	15,950	120
	Wood 2	15,970	120
	Globigerinoides sacculifer	16,480	120
	P. obliquiloculata	16,330	100
	P. obliquiloculata	16,760	110
	N. dutertrei	16,740	110
	Globorotalia tumida	16,290	110
	Mix benthics	17,690	130
	Mean planktics	16,520	150
	Δ B-P	1,170	220
1279–1285			
	P. obliquiloculata	16,900	130
	N. dutertrei	17,150	130
	Mix benthics	18,350	120
	Mean planktics	17,025	150
	Δ B-P	1,325	200
	Admiralty Island MD97-2138 1°S 146°E 1.9 km		
207–210			
	Globigerinoides sacculifer	18,780	120
	P. obliquiloculata	18,620	120
	N. dutertrei	18,270	120
	Globorotalia tumida	18,890	120
	N. dutertrei	18,620	120
	Mix benthics	20,590	140
	Mean planktics	18,636	150
	Δ B-P	1,954	220
211–215			
	Globigerinoides sacculifer	19,130	130
	P. obliquiloculata	18,970	130
	N. dutertrei	18,730	140
	Globorotalia tumida	18,960	120
	Mix benthics	20,590	150
	Mean planktics	18,950	200
	Δ B-P	1,640	250

References and Notes

1. K. Hughen *et al.*, *Science* **303**, 202 (2004).
2. R. Muscheler *et al.*, *Earth Planet. Sci. Lett.* **219**, 325 (2004).
3. W. S. Broecker, E. Clark, I. Hajdas, G. Bonani, *Paleoceanography* **19**, doi: 2003PA000974 (2004).
4. L. Beaufort, T. de Garidel-Thoron, A. C. Mix, N. G. Pisias, *Science* **293**, 2440 (2001).
5. Alex Wiedenhoeft of the Forest Products Laboratory in Madison, Wisconsin, advised us that it was not possible to identify milligram-sized fragments of tropical woods; rather, a hand-sized piece of wood would be required. He also commented that, once waterlogged, any wood would be dense enough to sink in seawater.
6. T. P. Guilderson, D. P. Schrag, M. A. Cane, *J. Clim.* **17**, 1147 (2004).
7. T. Guilderson, personal communication, 2004.
8. Because conventional radiocarbon ages are calculated using the Libby half life of 5568 years, we have used this value to calculate all age differences in this study for consistency.
9. E. Bard, *Paleoceanography* **3**, 635 (1988).
10. W. S. Broecker, W. C. Patzert, J. R. Toggweiler, M. Stuiver, *J. Geophys. Res.* **91**, 14345 (1986).
11. R. M. Key *et al.*, *Radiocarbon* **44**, 239 (2002).
12. J. F. Adkins, E. A. Boyle, *Paleoceanography* **12**, 337 (1997).
13. J. F. Adkins, K. McIntyre, D. P. Schrag, *Science* **298**, 1769 (2002).
14. K. Matsumoto, T. Oba, J. Lynch-Stieglitz, H. Yamamoto, *Quat. Sci. Rev.* **21**, 1693 (2002).
15. W. S. Broecker, T.-H. Peng, S. Trumbore, G. Bonani, W. Wolfli, *Global Biogeochem. Cycles* **4**, 103 (1990).
16. L. D. Keigwin, *J. Oceanogr.* **58**, 421 (2002).
17. K. Ohkushi, M. Uchida, N. Ahagon, T. Mishima, T. Kanematsu, *Nucl. Instrum. Methods Phys. Res. Sec. B,* **406**, 223 (2004).
18. S. J. Goldstein, D. W. Lea, S. Chakraborty, M. Kashgarian, M. T. Murrell, *Earth Planet. Sci. Lett.* **193**, 167 (2001).
19. M. Andree, *Radiocarbon* **29**, 169 (1987).
20. W. S. Broecker, K. Matsumoto, E. Clark, I. Hajdas, G. Bonani, *Paleoceanography* **14**, 431 (1999).
21. E. L. Sikes, C. R. Samson, T. P. Guilderson, W. R. Howard, *Nature* **405**, 555 (2000).
22. Four benthic-planktic age differences for LGM samples from cores off New Zealand are reported in (*21*): 2.12 thousand years (ky) (depth 2.07 km, planktic ^{14}C age 24.4 ky); 0.84 ky (depth 2.07 km, planktic ^{14}C age 25.6 ky); 3.48 ky (depth 2.7 km, planktic ^{14}C age 24.1 ky); and 0.75 ky (depth 1.3 km, planktic ^{14}C age 24.8 ky). We choose not to include them because they represent a different age range from those that we studied.
23. We thank L. Beaufort for providing the Admiralty core sample and T. Guilderson for providing a survey of coral-derived prebomb Δ^{14}C values. This paper is funded in part by a grant/cooperative agreement from the National Oceanic and Atmospheric Administration (NOAA). The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. Financial support was provided by NSF grant OCE 02-21979 and by NOAA Consortium on the Ocean's Role in Climate (CORC) grant UCSIO P.O. 10196097-003/NA17RJ1231. This study was funded in part by a Lamont-Doherty postdoctoral fellowship and a fellowship grant (award number CSEF CC3B) from the Comer Science and Education Foundation to S.B., Lamont-Doherty Earth Observatory contribution no. 6663.

2 July 2004; accepted 8 October 2004

Prospects for Building the Tree of Life from Large Sequence Databases

Amy C. Driskell,[1,2]* Cécile Ané,[1]†‡ J. Gordon Burleigh,[1]† Michelle M. McMahon,[1]† Brian C. O'Meara,[2]† Michael J. Sanderson[1]

We assess the phylogenetic potential of ~300,000 protein sequences sampled from Swiss-Prot and GenBank. Although only a small subset of these data was potentially phylogenetically informative, this subset retained a substantial fraction of the original taxonomic diversity. Sampling biases in the databases necessitate building phylogenetic data sets that have large numbers of missing entries. However, an analysis of two "supermatrices" suggests that even data sets with as much as 92% missing data can provide insights into broad sections of the tree of life.

More than 100,000 species—about 6% of all those known to science—have at least one molecular sequence archived in public databases, but what fraction of these sequences is phylogenetically informative? Here, we examine two large samples of proteins and show how the answer depends on the pattern of homology among sequences and the distribution of sequences among taxa. We then parse these databases into phylogenetic supermatrices for metazoans and green plants. Although the databases have sampling biases that cause these matrices to be very sparse, they can still provide useful information for building the tree of life.

We examined the phylogenetic information content of the Swiss-Prot database of 120,000 sequences for nearly 7500 taxa and a "taxonomically enriched" subset of GenBank, which consisted of 185,000 amino acid sequences for more than 16,000 green plant taxa (*1*). Clusters of putative homologs were identified via $N \times N$ BLAST searches (*2*). Clustering procedures involve trade-offs among the reliability of homology assessment, the taxonomic breadth, and the accuracy of tree inference. The trade-offs are controlled by the stringency of homology searches and can be adjusted to maximize the phylogenetic utility of resulting clusters, on the basis of the depth and breadth of the phylogenetic question to be addressed (*3*, *4*). Clusters containing at least four taxa are termed "minimal phylogenetic clusters," because unrooted trees with fewer than four taxa contain no information about relations. Although minimal phylogenetic clusters were a small fraction of all clusters found [6.5% and 2.3% for Swiss-Prot and GenBank, respectively (Table 1)], they retained about one-third of the original sequences and a substantial fraction (74% and 95%) of the taxonomic diversity originally contained in the sample.

We screened minimal phylogenetic clusters for the presence of paralogs with a phylogenetic test of orthology (*5*). A species tree cannot be easily deduced from a cluster containing both orthologs and paralogs, although methods for this have been proposed (*6*–*8*). Screening reduced the candidate minimal clusters to smaller sets of orthologous "single-copy" clusters retaining only 24% and 21% of the original sequences, but still covering 59% and 89% of the original taxa in Swiss-Prot and GenBank, respectively (Table 1). These sequences are very sparsely distributed among taxa as measured by their "densities" (Table 1) (*9*).

Further assessment of the phylogenetic utility of these data requires consideration of how the data should be parsed for phylogenetic analyses. One approach is to build gene trees from individual clusters and to assemble these trees using supertree methods (*3*, *10*). Supertree methods require at least partial taxonomic overlap between trees. A set of trees (each inferred from a cluster) with enough taxonomic overlap to allow supertree construction is a "grove" (*1*). The minimum number of groves in a database is a lower bound on the number of supertrees required to encompass all its sequence data. The single-copy green plant proteins form at least 15 groves (Table 1). The largest of these groves minimally includes trees from 814 clusters and contains more than 14,000 taxa—87% of all the green plant taxa in the GenBank database. Swiss-Prot has at least eight times as many groves, which reflects its greater taxonomic breadth but higher fragmentation (Table 1). Both data sets also contain a small number of "orphans," clusters with no taxonomic overlap with other clusters.

[1]Section of Evolution and Ecology, [2]Center for Population Biology, University of California, One Shields Avenue, Davis, CA 95616, USA.

*To whom correspondence should be addressed. E-mail: acdriskell@ucdavis.edu

†These authors contributed equally to this work.

‡Present address: Department of Statistics, University of Wisconsin, Medical Science Center, 1300 University Avenue, Madison, WI 53706, USA.

A second and more widely used strategy is to concatenate clusters into multigene (protein) data matrices. Increasing the sequence data per taxon should improve accuracy (*11*, *12*), a theoretical result supported by several recent phylogenomic studies (*13*–*15*). However, as with supertree construction, the taxonomic structure of the set of clusters limits the size of data matrices that can be assembled from them. To explore these limits, we used an exact algorithm derived in the context of a well-known graph problem on "bicliques" (*5*, *16*) to enumerate all possible concatenated matrices that are both maximal (not contained in larger matrices) and complete (no missing sequences). The largest complete multiprotein Swiss-Prot and GenBank matrices have either many taxa and few proteins or the reverse; none has large numbers of proteins and taxa simultaneously (Fig. 1), and these matrices retain only a small fraction of the original taxonomic diversity in the databases (Table 1). The numbers of taxa and genes in a matrix can be greatly increased by allowing missing sequences, or "holes," in the matrix (*17*, *18*). Such "supermatrices" are also affected by the grove structure of the database, because the missing entries induce patterns of partial overlap among taxa.

Statistics on clusters, groves, and bicliques offer a glimpse of the potential phylogenetic information content of the sequence databases. To explore the feasibility of realizing this potential, we assembled supermatrices spanning a small (but broad) taxonomic sample from each database. We identified the largest grove and constructed a supermatrix from complete matrices having at least 10 clusters and four taxa (*1*). The green plant matrix retained representatives across the group, but Swiss-Prot was culled to metazoans (plus fungal outgroups) to avoid conflicting gene histories of nuclear, mitochondrial, and chloroplast genomes. The resulting two supermatrices had comparable numbers of taxa and shared no sequences. The Swiss-Prot supermatrix had 70 taxa × 1131 genes (6623 sequences and 469,497 characters), an average of 95 genes per taxon. The green plant supermatrix contained 69 taxa × 254 genes (2777 sequences and 96,698 characters), an average of 40 genes per taxon. These are among the largest supermatrices yet analyzed for phylogenetic inference, as well as the sparsest, with 92% and 84% missing entries, respectively (Fig. 2). Yet, although sparse, they are 46 and 75 times as dense, respectively, as the single-copy minimal sequence collections from which they were derived (Table 1).

Trees from these matrices (*1*) broadly agree with conventional views on phylogenetic relations, and many nodes, particularly in the Swiss-Prot tree, are well supported by bootstrap values (figs. S1 and S2). However, some of the "backbone" of the green plant tree is weakly supported, and some unconventional and likely incorrect relations are depicted in both topologies. Many of these unconventional relations have been seen in previous molecular studies (*19*–*22*), including even the nonmonophyly of monocots (*23*). To characterize the sources of these signals, we compared (Fig. 3) every clade in each individual protein tree against the final trees (figs. S1 and S2). Compared with the green plant supermatrix, fewer of the proteins in the Swiss-Prot data conflict with the relations in the final tree. The level of conflict in green plants is remarkably high: More individual protein trees conflict with any given final clade than support it. Nonetheless, many nodes on this tree are still well supported by bootstrap values. Surprisingly little relation was seen between the number of proteins supporting a node and its bootstrap score. For example, placement of the lepidopteran *Spodoptera frugiperda* within the Diptera receives 96% bootstrap support, but only one protein cluster of the 1131 sampled contains sequence for both *Spodoptera* and other dipterans (none contain

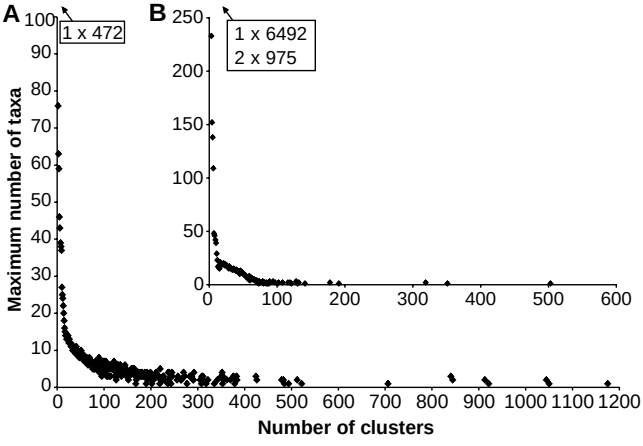


Fig. 1. Size distribution of the maximal complete supermatrices (maximal bicliques) for two data sets. Each point represents the biclique with the largest number of taxa for a particular number of clusters. (**A**) Swiss-Prot. (**B**) Green plants from GenBank. Bicliques on the tails are not shown, but their sizes are indicated in the inset boxes.

Table 1. Summary statistics on the phylogenetic information content of proteins from Swiss-Prot and GenBank (green plant only) databases. "Minimal phylogenetic clusters" contain at least four taxa. "Nontrivial" bicliques contain at least two genes and four taxa. Density = (number of cells in the taxon × cluster matrix containing ≥ 1 sequence)/(number of taxa × number of clusters) (*1*). See text for other definitions.

Database statistics	Swiss-Prot release 40.29	GenBank release 137 (green plant)
Summary statistics		
Number of sequences in release	121,218	185,418
Total number of sequences clustered	121,218	185,089
Total number of taxa	7449	16,348
Total number of clusters	64,712	59,144
Minimal phylogenetic clusters		
Number of clusters	4214 (6.5%)	1365 (2.3%)
Sequence coverage	41,812 (34%)	65,113 (35%)
Taxon coverage	5538 (74%)	15,599 (95%)
Single-copy clusters		
Number of clusters	3592 (5.6%)	853 (1.4%)
Sequence coverage	28,742 (24%)	39,443 (21%)
Taxon coverage	4404 (59%)	14,502 (89%)
Density	0.0018	0.0021
Groves		
Minimum number of groves	123	15
Number of orphan clusters	67	7
Minimum number of clusters in largest grove	3183	814
Minimum number of sequences in largest grove	25,272 (21%)	38,700 (49%)
Minimum number of taxa in largest grove	2695 (36%)	14,169 (87%)
Bicliques		
Number of nontrivial maximal bicliques	43,576	5587
Sequence coverage	23,855 (20%)	15,092 (8.2%)
Taxon coverage	1449 (19%)	4230 (26%)
Number of clusters in biclique set	3187 (4.9%)	645 (1.1%)
Largest biclique (in terms of taxa)	2 × 76	2 × 975
Largest biclique (in terms of clusters)	352 × 4	70 × 4

sequence from *Spodoptera* and both *Anopheles* species) and can therefore shed light on the final topology. Other relations, such as the placement of *Spathiphyllum* outside of the main clade of monocots, are not supported by a single protein and must therefore be an emergent property of weak signals buried within several proteins.

These supermatrices differ from other recent phylogenomic analyses of diverse taxa (*13*–*15*, *18*) in two important respects. First, the taxa and proteins in our supermatrices were determined largely by the structure of the databases rather than by decisions of the investigators. Second, they contain more taxa (two or more times as many), but one-fourth or fewer proteins per taxon, with a substantially lower overall density. In two studies with <15 taxa, bootstrap support was uniformly very high across the tree (*14*, *15*) and could be obtained even with a subset of sequences (*14*). In two other studies (*13*, *18*) with larger taxon sets (30 to 36 taxa), many clades were strongly supported but some were not—an expected consequence of increased taxon sampling (*24*). Our results support the general conclusion of these other studies that combining many genes can provide strong support for nodes in a large and complex tree, with the added twist that genes need not be sampled evenly across taxa and a surprising amount of missing data is tolerable (*18*, *25*). Nevertheless, even our threshold of 10 proteins per taxon was not enough to prevent recovering some unconventional results, as our surprising findings regarding monocots document. Such results, if they stem from genomewide long branch–attraction artifacts (*26*), will also occur in phylogenomic studies unless problematic taxa are deliberately or accidentally excluded and will only be overcome by additional taxon sampling or intensive analytical efforts. Therefore, we conclude that exploitation of existing databases, taking into account the inherent sample biases of the data, provides a cost-effective complement to intensive genomewide sequencing efforts, especially if we wish to include large numbers of taxa or remote corners of the tree of life.

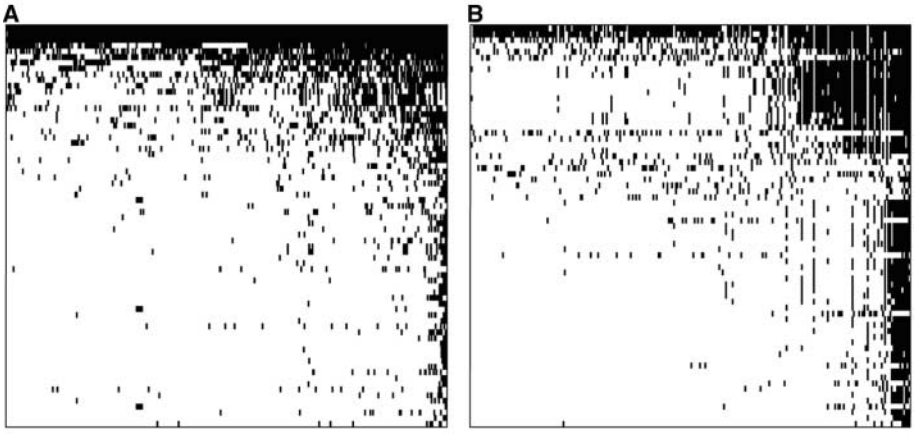


Fig. 2. Visualization of the distribution of sequences among taxa for the supermatrices. (**A**) SwissProt metazoan supermatrix. (**B**) GenBank green plant supermatrix. Columns correspond to clusters (proteins); rows to taxa, ordered so that density increases to upper right. Black indicates a sequence is present in that cluster for that taxon.

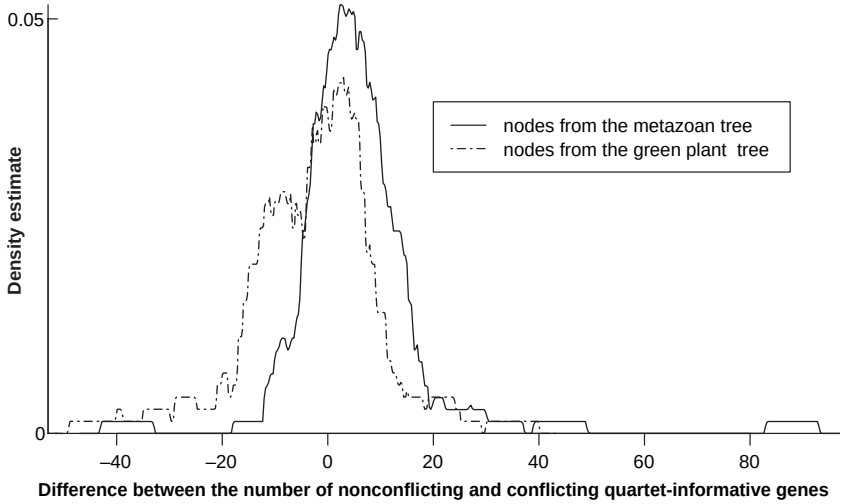


Fig. 3. Protein-by-protein distribution of support for clades in supermatrix trees (figs. S1 and S2). A protein tree is "quartet informative" for a given branch (bipartition) of the supermatrix tree if its data set had sequence for at least one taxon in each of the four groups attached to this branch. It is considered "nonconflicting" if it has at least one most-parsimonious (MP) tree displaying the bipartition (pruned to the taxon set for the protein) or if all of its MP trees display a polytomy in place of the branch. Otherwise, the protein is classified as "conflicting." Plots are smoothed histograms (density estimates) of the number of quartet-informative proteins that are nonconflicting minus the number that conflict. Values to the right of the origin are therefore branches in which the majority of proteins do not conflict. Density estimation was done with *R* statistical software (*27*), using a rectangular kernel and bandwidth of 3 for both trees.

References and Notes

1. Materials and methods are available as supporting material on *Science* Online.
2. I. Dondoshansky, *BLASTCLUST vers. 6.1*, (National Center for Biotechnology Information, Bethesda, MD, 2002).
3. O. R. P. Bininda-Emonds, S. G. Brady, J. Kim, M. J. Sanderson, *Pac. Symp. Biocomp.* **6**, 547 (2001).
4. Z. Yang, *Syst. Biol.* **47**, 125 (1998).
5. M. J. Sanderson, A. C. Driskell, R. H. Ree, O. Eulenstein, S. Langley, *Mol. Biol. Evol.* **20**, 1036 (2003).
6. R. D. M. Page, *Mol. Phylogenet. Evol.* **14**, 89 (2000).
7. J. Kim, B. Salisbury, *Pac. Symp. Biocomp.* **6**, 571 (2001).
8. L. Arvestad, A.-C. Berglund, J. Lagergren, B. Sennblad, *Bioinformatics* **19** (suppl. 1), i7 (2003).
9. M. J. Sanderson, A. C. Driskell, *Trends Plant Sci.* **8**, 374 (2003).
10. M. J. Sanderson, A. Purvis, C. Henze, *Trends Ecol. Evol.* **13**, 105 (1998).
11. P. L. Erdös, M. A. Steel, L. A. Székely, T. J. Warnow, *Rand. Struct. Algorithms* **14**, 153 (1999).
12. D. M. Hillis, *Nature* **383**, 130 (1996).
13. E. Bapteste *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1414 (2002).
14. A. Rokas, B. Williams, N. King, S. Carroll, *Nature* **425**, 798 (2003).
15. E. Lerat, V. Daubin, N. Moran, *PLoS Biol.* **1**, 101 (2003).
16. G. Alexe *et al.*, *DIMACS Tech. Rep. 2002-52* (2002).
17. C. Yan, J. G. Burleigh, O. Eulenstein, in preparation.
18. H. Philippe *et al.*, *Mol. Biol. Evol.* **21**, 1740 (2004).
19. A. M. D'Erchia, C. Gissi, G. Pesole, C. Saccone, U. Arnason, *Nature* **381**, 597 (1996).
20. M. J. Phillips, D. Penny, *Mol. Phylogenet. Evol.* **28**, 171 (2003).
21. R. Zardoya, Y. Cao, M. Hasegawa, A. Meyer, *Mol. Biol. Evol.* **15**, 506 (1998).
22. M. R. Duvall, A. B. Ervin, *Mol. Phylogenet. Evol.* **30**, 97 (2004).
23. D. E. Soltis *et al.*, *Ann. Mo. Bot. Gard.* **84**, 1 (1997).
24. J. Kim, *Syst. Biol.* **45**, 363 (1996).
25. J. J. Wiens, *Syst. Biol.* **52**, 528 (2003).
26. J. Felsenstein, *Biol. J. Linn. Soc.* **16**, 183 (1978).
27. See www.r-project.org/.
28. This research was supported by the National Science Foundation. Thanks to R. Piaggio and O. Eulenstein for insights on the properties of groves and Wen-Chieh Chang for biclique code. We also thank D. Fernández-Baca and J. Kim.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1172/DC1
Materials and Methods
Figs. S1 and S2

28 June 2004; accepted 14 September 2004

Realistic Species Losses Disproportionately Reduce Grassland Resistance to Biological Invaders

Erika S. Zavaleta* and Kristin B. Hulvey

Consequences of progressive biodiversity declines depend on the functional roles of individual species and the order in which species are lost. Most studies of the biodiversity–ecosystem functioning relation tackle only the first of these factors. We used observed variation in grassland diversity to design an experimental test of how realistic species losses affect invasion resistance. Because entire plant functional groups disappeared faster than expected by chance, resistance declined dramatically with progressive species losses. Realistic biodiversity losses, even of rare species, can thus affect ecosystem processes far more than indicated by randomized-loss experiments.

Rapid biodiversity losses at both local and global scales disproportionately involve species with particular values of traits such as size, trophic position, rarity, distribution, and degree of specialization (*1*–*5*). While the evolutionary consequences of the nonrandomness of biodiversity loss have been explored recently (*6*), our understanding of the functional consequences of nonrandom loss remains limited mainly to studies of the declines of individual species (*7*, *8*) [but see (*9*)]. These studies focus on keystone species or on species that were once widespread and abundant. However, most species in ecosystems—and most at-risk species—are less dominant and exert what influence they have on ecosystem functioning in relatively small numbers (*10*–*12*).

The prevailing experimental approach to testing the functional consequences of biodiversity change isolates the effects of species or functional group richness by (i) randomizing species composition and (ii) equalizing species abundances within each richness treatment (*13*–*16*). This approach provides little information about the importance of species loss order ["community disassembly" (*5*)] or relative species abundances for ecosystem processes. Recent studies generating nonrandom species losses through removals of rare and uncommon species from natural communities provide more insight about the effects of nonrandom loss on ecosystem functioning (*17*, *18*), but local abundance or rarity alone is an imperfect predictor of species loss order (*19*) (Table 1).

Patterns of nonrandom change and variation in diversity can be better assessed with tools specifically designed for the measurement of extinction order, such as nested subset analysis (*20*, *21*). Nested subset analysis quantifies the degree to which species disappearances are ordered, most often across a series of habitat fragments of decreasing size or increasing isolation (*22*, *23*). Although seldom extended in this way, nested subset analysis can also be used to quantify the order in which species are lost through space or time. It bases this order solely on observations of species occurrences across sites or dates rather than on assumptions about specific drivers of diversity loss or species characteristics thought to be associated with species vulnerability to local extinction (*20*, *24*). Nested subset analysis can thus provide an empirical species loss order through either space or time at a specified scale and for a particular community type.

We observed 4 years of spatial variation in the plant species richness of plots in a California grassland ecosystem to quantify the degree to which changes in richness occurred in a consistent, nested order (*25*). We treated nested patterns of species occurrences through space as indicators of a representative order of local extinction, because similar mechanisms appear to affect richness through space and time in our study system. Variation through space in species richness at our level, treeless site appears to reflect heterogeneous soil conditions resulting from gopher activity (up to present) and past mechanical soil disturbance and vehicular compaction (as recently as 30 years ago) (*26*). The resulting heterogeneity likely affects species richness both directly and indirectly through effects on productivity (*27*–*29*) and invasibility (*30*, *31*). These mechanisms resemble important drivers of grassland biodiversity change in California and elsewhere, including soil disturbances associated with agriculture and livestock production, biological invasions, and increased productivity due to anthropogenic nitrogen fertilization (*32*, *33*),

Table 1. Species compositions and traits comprising each experimental diversity level. Values are target numbers of stems for each species in a given treatment. Life history traits: early annual (E), indeterminate annual/late annual (L), biennial/perennial (P). Functional types were defined as early grasses, early forbs, nitrogen fixers, late forbs, and perennials based on groupings from previous California grassland studies (*13, 38*). Abundance ranks are means based on field observations of neighborhood relative abundances across plots of varying richness; 1 indicates most locally abundant species.

Species	Diversity level						Type	Life history	N fixer	Abundance rank
No. species	3	6	9	12	15	21				
No. functional types	2	3	3	4	4	5				
Avena barbata	206	149	110	98	74	24	Grass	E		1
Bromus hordeaceus	41	41	41	41	41	41	Grass	E		2
Geranium dissectum	21	20	21	20	21	21	Forb	E		6
Bromus diandrus		13	13	12	12	12	Grass	E		7
Lolium multiflorum		35	28	24	18	7	Grass	E		5
Vicia sativa		10	10	10	10	10	Forb	L	yes	8
Avena fatua			34	33	34	33	Grass	E		4
Erodium botrys			9	9	9	9	Forb	E		9
Vulpia microstachys			2	2	2	2	Grass	E		17
Anagallis arvensis				10	27	63	Forb	L		3
Briza minor				7	10	15	Grass	E		10
Epilobium brachycarpum				2	3	5	Forb	L		14
Crepis vesicaria					2	2	Forb	P		19
Torilis arvensis					3	3	Forb	L		13
Trifolium hirtum					2	2	Forb	L	yes	18
Danthonia californica						5	Grass	P		12
Hemizonia congesta						2	Forb	L		16
Hordeum murinum						2	Grass	E		15
Medicago polymorpha						2	Forb	L	yes	21
Phalaris aquatica						2	Grass	P		20
Rumex acetosella						6	Forb	P		11

Environmental Studies Department, University of California, Santa Cruz, CA 95064, USA.

*To whom correspondence should be addressed. E-mail: zavaleta@ucsc.edu

although they might capture less well species loss orders likely to be associated with other drivers of grassland biodiversity change, such as woody encroachment and climate change.

Species-by-site occurrence matrices at our study location were significantly nested in all years ($T = 19.4°$ to $23.4°$, $P \ll 0.001$) (*34*). Local rarity was an inconsistent predictor of loss order, with some abundant species absent from low-richness plots and some uncommon species present in plots of both high and low richness (table S1). The most abundant species in our richest plots, *Anagallis arvensis* (Primulaceae), seldom occurred at all in plots below the median richness level (Table 1). All of the 38 species at our site maintained their approximate ranked positions in the nested order of diversity changes across years as individual plots increased or declined in species richness (table S1). This indicated that the order of species losses and gains through space from grassland patches is robust to interannual variability and change at our site. We used this observed nested order of species loss and gain to design a test of how these ordered changes in species richness influence grassland resistance to invasion, an ecosystem function of growing conservation relevance as invasions accelerate (*35*).

We constructed outdoor microcosm communities by planting locally collected seeds at six levels of species richness to reflect (i) the observed, nested order of local-scale variation in species richness and (ii) variation in relative species abundances that accompanied variation in species richness at our site (Table 1). At all species richness levels, species abundance ranks followed a log series–like distribution (*11*), with abundances varying by two orders of magnitude and most species rare as observed in our field plots (Fig. 1). Constructed communities differed in the degree of dominance (dominance was higher in more species-poor treatments) but not the identity of the most abundant species. This allowed us to distinguish the effects of progressive changes in species richness from effects associated with the identity of the dominant species (*17*, *36*). We invaded half of the communities at each richness level with *Centaurea soltitialis* L. (yellow starthistle), an expanding California grassland exotic that has already caused considerable ecological and economic damage (*37*).

Starthistle biomass increased with progressive species loss by more than 100% from the most diverse to the least diverse communities ($F = 5.58$, $P = 0.002$, $R^2 = 0.59$) (Fig. 2). Starthistle production and flower number were strongly correlated (Pearson coefficient = 0.790, Bartlett's $\chi^2 = 33.8$, $P < 0.001$), such that starthistle reproduction also increased with declining richness. Lost species effects on invader performance were highly disproportionate to their abundance in the community. For example, mean starthistle biomass was >60% lower in 20-species than in 15-species communities, even though the additional 5 species in the 20-species communities together comprised <4% of resident production and <3% of the total number of stems. Similarly, mean starthistle biomass was >70% lower in 20-species than in 12-species communities, while the 8 additional species made up <5% of resident production and <7% of total stems.

Our findings differ from the results of studies at the same site testing the effects of randomized changes in species richness on starthistle (*38*). In this earlier work, no differences in invader production occurred among diversity treatments with >1 species. Functional group number was held constant across diversity levels with >1 species in the earlier study. Our nested subset analyses show, however, that in our study system, progressive diversity declines are not evenly distributed across species varying in key functional traits (Table 1). In our system, species-poor assemblages contain mainly early annuals, with no or few indeterminate, late-season, nitrogen-fixing, or perennial species that overlap strongly with starthistle in the location or timing of resource uptake. Progressively richer communities gain late-season annual forbs and perennial species, the former of which were shown to compete effectively against starthistle (which is also a late-season annual forb) in monoculture (*38*). The effects of reducing species richness while holding functional group number constant, while theoretically important, thus do not represent the effects of real biodiversity variation. Consequently, the incorporation of realistic species loss order into our experimental design profoundly altered the observed relation between diversity and invasibility.

The decline in functional diversity that accompanies realistic species losses can also influence the degree to which invasions affect ecosystem processes. Starthistle increased total biomass (resident + starthistle) more at low than at high resident species richness levels ($F = 6.77$, $P = 0.01$, $R^2_{adj} = 0.17$), reflecting greater total starthistle biomass in the more species-poor communities. The per-unit impact of starthistle on resident biomass, in contrast, increased with species richness, from <0.3 to nearly 1 g resident biomass lost per g starthistle present ($F = 4.79$, $P = 0.04$) (Fig. 3). Most likely,

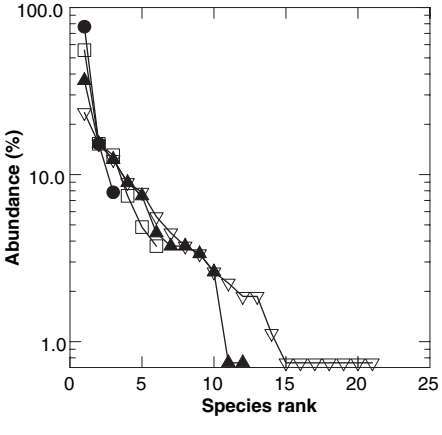


Fig. 1. Target rank-abundance curves for 3- (closed circles), 6- (open squares), 12- (closed triangles), and 21-species (open triangles) treatments. Target treatment compositions display log series–like patterns of relative abundance, with most species relatively uncommon and a single dominant species across all diversity levels.

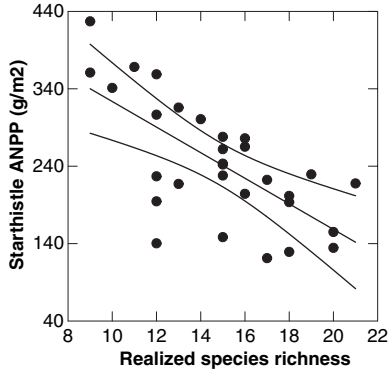


Fig. 2. The relation between realized species richness and starthistle aboveground net primary production (ANPP). Realized richness values reflect effects of negative (species that failed to establish) and positive (unplanted volunteers) deviations from treatment richness levels. Target diversity $F = 5.58$, $P = 0.002$, $R^2 = 0.59$; realized deviations $F = 7.85$, $P = 0.01$. 95% confidence limits are shown.

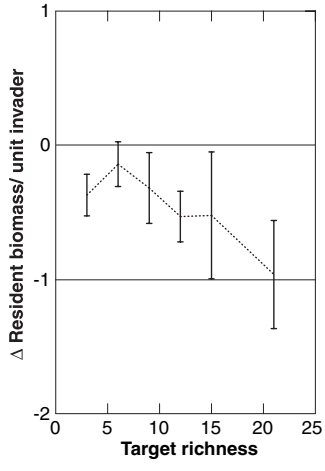


Fig. 3. Effect of species richness on per-unit invader impact (change in resident biomass/gram of starthistle biomass) ± 1 SE. This metric = 0 when resident biomass is unaffected by starthistle invasion and –1 when 1 g of resident biomass is displaced by each g of establishing starthistle. $N = 5$, $F = 4.79$, $P = 0.04$, $R^2_{adj} = 0.12$.

this reflects more complete use of certain resources by the more species-rich assemblages. As a result, starthistle added substantial biomass to species-poor communities while mainly displacing resident biomass in species-rich communities. Invasibility can thus decline while per-unit invader impact on the resident community increases, underscoring the importance of measuring both.

This study helps bridge the gap between our understanding of general biodiversity-function relations and the role of extinction order in determining the consequences of biodiversity loss. Additional experiments are needed to assess the consequences of ordered species losses for other ecosystems and ecosystem functions, as well as to expand research designs to incorporate species losses occurring through time at larger spatial scales. If, as we found, important functional traits disappear more rapidly than expected by chance in other communities, the ecosystem consequences of real biodiversity losses—even of rare species—will often exceed expectations based on randomized diversity studies.

References and Notes

1. S. P. Lawler, J. J. Armesto, P. Kareiva, in *The Functional Consequences of Biodiversity*, A. P. Kinzig, S. W. Pacala, D. Tilman, Eds. (Princeton Univ. Press, Princeton, NJ, 2001), vol. 33, pp. 294–313.
2. K. Henle, K. F. Davies, M. Kleyer, C. Margules, J. Settele, *Biodivers. Conserv.* **13**, 207 (2004).
3. L. R. Belyea, J. Lancaster, *Oikos* **86**, 402 (1999).
4. K. A. McDonald, J. H. Brown, *Conserv. Biol.* **6**, 409 (1992).
5. B. Fox, *Evol. Ecol.* **1**, 201 (1987).
6. A. Purvis, P. M. Agapow, J. L. Gittleman, G. M. Mace, *Science* **288**, 328 (2000).
7. J. A. Estes, M. T. Tinker, T. M. Williams, D. F. Doak, *Science* **282**, 473 (1998).
8. D. E. Blockstein, *Science* **279**, 1831c (1998).
9. R. S. Ostfeld, K. LoGiudice, *Ecology* **84**, 1421 (2003).
10. F. W. Preston, *Ecology* **43**, 185 (1962).
11. A. E. Magurran, *Ecological Diversity and Its Measurement* (Princeton Univ. Press, Princeton, NJ, 1988).
12. International Union for the Conservation of Nature (IUCN), (IUCN, Gland, Switzerland, 2003), vol. 2004.
13. D. Hooper, P. Vitousek, *Science* **277**, 1302 (1997).
14. D. Tilman *et al.*, *Science* **277**, 1300 (1997).
15. S. Naeem, S. Li, *Nature* **390**, 507 (1997).
16. M. Loreau, S. Naeem, P. Inchausti, Eds., *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (Oxford Univ. Press, New York, 2002).
17. M. D. Smith, A. K. Knapp, *Ecol. Lett* **6**, 509 (2003).
18. K. G. Lyons, M. W. Schwartz, *Ecol. Lett.* **4**, 358 (2001).
19. D. Rabinowitz, in *The Biological Aspects of Rare Plant Conservation*, H. Synge, Ed. (Wiley, London, 1981), pp. 205–217.
20. D. H. Wright, J. H. Reeves, *Oecologia* **92**, 416 (1992).
21. D. H. Wright, B. D. Patterson, G. M. Mikkelson, A. Cutler, W. Atmar, *Oecologia* **113**, 1 (1998).
22. B. D. Patterson, W. Atmar, *Biol. J. Linn. Soc.* **28**, 65 (1986).
23. R. Kadmon, *Ecology* **76**, 458 (1995).
24. E. A. Hadly, B. A. Maurer, *Evol. Ecol. Res.* **3**, 477 (2001).
25. Materials and methods are available as supporting material on *Science* Online.
26. R. A. Hobbs, H. A. Mooney, *Ecology* **72**, 59 (1991).
27. E. Weiher, S. Forbes, T. Schauwecker, J. B. Grace, *Oikos* **106**, 151 (2004).
28. B. L. Foster, K. L. Gross, *Ecology* **79**, 2593 (1998).
29. D. E. Goldberg, T. E. Miller, *Ecology* **71**, 213 (1990).
30. L. F. Huenneke, S. P. Hamburg, R. Koide, H. A. Mooney, P. M. Vitousek, *Ecology* **71**, 478 (1990).
31. E. I. Newman, *Nature* **244**, 310 (1973).
32. A. R. Watkinson, S. J. Ormerod, *J. Appl. Ecol.* **38**, 233 (2001).
33. C. J. Stevens, N. B. Dise, J. O. Mountford, D. J. Gowing, *Science* **303**, 1876 (2004).
34. W. Atmar, B. D. Patterson, *Nestedness Temperature Calculator* (AICS Research, Inc., University Park, NM, and The Field Museum, Chicago, IL, 1995); www.aics-research.com/nestedness/tempcalc.html.
35. P. M. Vitousek, C. M. D'Antonio, L. L. Loope, M. Rejmanek, R. Westbrooks, *N. Z. J. Ecol.* **21**, 1 (1997).
36. L. W. Aarssen, *Oikos* **80**, 183 (1997).
37. F. S. Chapin *et al.*, *Nature* **405**, 234 (2000).
38. J. S. Dukes, *Oecologia* **126**, 563 (2001).

39. We thank N. Chiariello, C. Field, T. Tobeck, E. Cleland, E. Hadly, J. Kriewall, D. Croll, R. Shaw, the Carnegie Institution of Washington, and the Jasper Ridge Biological Preserve for their valuable contributions and assistance. K. Andonian, D. Doak, J. Dukes, G. Gilbert, P. Holloran, D. Hooper, K. Honey, D. Letourneau, J. Levine, B. Tershy, and two anonymous reviewers improved the manuscript. This project was generously supported by a David H. Smith Conservation Research Fellowship through The Nature Conservancy.

Supporting Online Materials

www.sciencemag.org/cgi/content/full/306/5699/1175/DC1
Materials and Methods
Table S1
References

12 July 2004; accepted 23 September 2004

Extinction and Ecosystem Function in the Marine Benthos

Martin Solan,[1]* Bradley J. Cardinale,[2] Amy L. Downing,[3] Katharina A. M. Engelhardt,[4] Jennifer L. Ruesink,[5] Diane S. Srivastava[6]†

Rapid changes in biodiversity are occurring globally, yet the ecological impacts of diversity loss are poorly understood. Here we use data from marine invertebrate communities to parameterize models that predict how extinctions will affect sediment bioturbation, a process vital to the persistence of aquatic communities. We show that species extinction is generally expected to reduce bioturbation, but the magnitude of reduction depends on how the functional traits of individual species covary with their risk of extinction. As a result, the particular cause of extinction and the order in which species are lost ultimately govern the ecosystem-level consequences of biodiversity loss.

Marine coastal ecosystems are among the most productive and diverse communities on Earth (*1*) and are of global importance to climate, nutrient budgets, and primary productivity (*2*). Yet, the contributions that coastal ecosystems make to these ecological processes are compromised by human-induced stresses, including overfishing, habitat destruction, and pollution (*3*–*5*). These stressors particularly impact benthic (bottom-living) invertebrate communities because many species are sedentary and cannot avoid disturbance. Thus, marine coastal ecosystems are likely to experience a large proportional change in biodiversity should present trends in human activity continue (*6*–*8*).

Given these prospects, researchers have recently asked how the loss of biodiversity might alter the functioning of marine coastal ecosystems. Like most studies to date, these experiments have manipulated diversity by assembling random subsets of species drawn from a common pool of taxa (*9*–*11*). This approach (*12*, *13*) may be useful for understanding the theoretical consequences of diversity loss but is unrealistic in the sense that it assumes species can go extinct in any order. Extinction, however, is generally a nonrandom process (*14*) with risk determined by life-history traits such as rarity, body size, and sensitivity to environmental stressors like pollution (*15*–*18*). Interspecific differences in extinction risk have implications for the ensuing changes in trophic interactions and community structure (*18*, *19*), such that the ecosystem-level consequences of random versus ordered extinctions are likely to be fundamentally different (*14*, *20*–*22*).

Here we explore how various scenarios of extinction for marine benthic invertebrates are likely to influence bioturbation (the biogenic mixing of sediment)—a primary determinant of sediment oxygen concentrations which, in turn, influences the biomass of organisms, the rate of organic matter decomposition, and the regeneration of nutrients vital for primary productivity (*23*, *24*).

[1]Oceanlab, University of Aberdeen, Main Street, Newburgh, Aberdeenshire, Scotland AB41 6AA. [2]Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA 93106, USA. [3]Department of Zoology, Ohio Wesleyan University, Delaware, OH 43015, USA. [4]University of Maryland Center for Environmental Science, Appalachian Laboratory, 301 Braddock Road, Frostburg, MD 21532–2307, USA. [5]Department of Biology, University of Washington, Box 351800, Seattle, WA 98195, USA. [6]Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia, Canada V6T 1Z4.

*To whom correspondence should be addressed. E-mail: m.solan@abdn.ac.uk
†All authors contributed equally to this work.

Using a comprehensive study of 139 benthic invertebrate species that inhabit Inner Galway Bay, Ireland (*25*), we parameterized models that predict how species extinction is likely to affect the biogenic mixing depth (BMD), an indicator of bioturbation that can be measured from sediment profile images (Fig. 1). To estimate species contributions to the BMD, we used an index of bioturbation potential (BP_i, Equation S1) that accounts for each species' body size, abundance, mobility, and mode of sediment mixing. We used data from monthly samples (over 1 year) of the benthic community to empirically derive a relation (Equation S2) between the BMD and the bioturbation potential of the community (BP_c). Using this relation, we performed numerical simulations to explore how the BMD is expected to change when species go extinct at random versus ordered by their sensitivities to environmental stress, body size, or population size (*25*). As the functional consequences of extinction are known to depend on the response of surviving species (*19*, *20*, *26*), we simulated two different types of community interactions (*8*). First, we used a model in which species do not interact with one another; thus, surviving species do not exhibit compensatory responses (changes in population size) after extinction. This scenario leads to complete loss of bioturbation performed by an extinct species and represents a "worst-case" scenario. Second, we used an interactive model of community assembly in which species' abundances are limited by competition with other members of their functional guild (i.e., species with similar bioturbation modes but not necessarily similar extinction risks). This represents a "best-case" scenario that assumes compensation is additive and substitutions of abundance maintain total community density [i.e., full numerical compensation (*25*)].

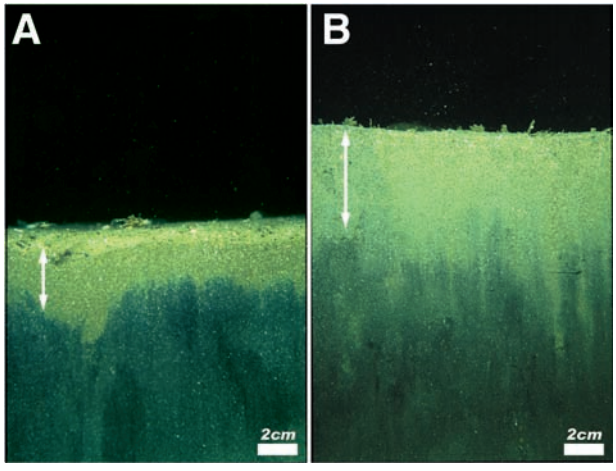


Fig. 1. The biogenic mixing depth (BMD, white arrows) of sediments [(**A**), site 1; (**B**), site 2] in Inner Galway Bay, Ireland. BMD was related to the bioturbation potential of a community (BP_c), an index that accounts for each species' population size and life-history traits (body size, mobility, mode of bioturbation) to estimate the capacity of a community to mix sediments (*25*).

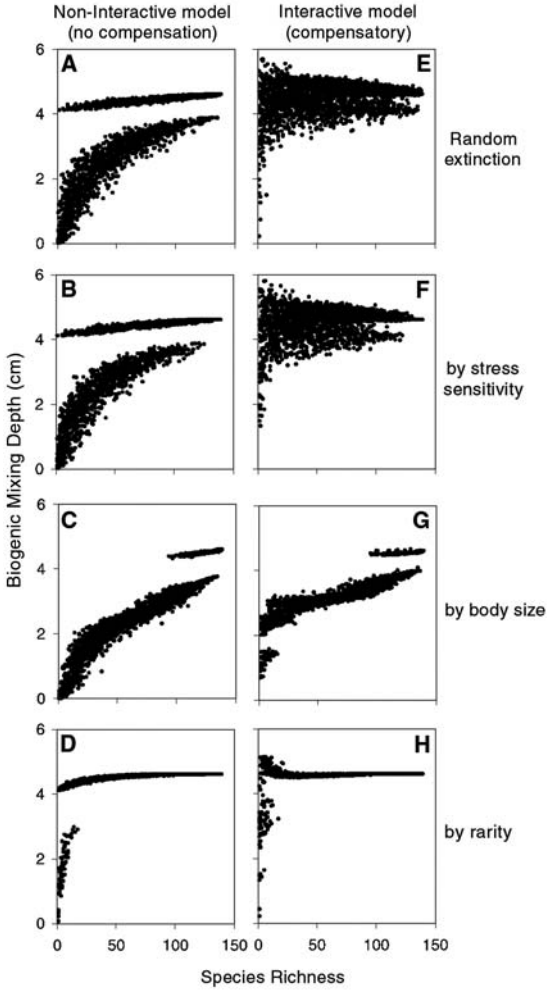


Fig. 2. Predicted changes in the BMD following benthic invertebrate extinctions. Each panel shows the results of 20 simulations per level of species richness, constrained by a probabilistic order of species extinction (indicated on the right). Simulations (**A**), (**B**), (**C**), and (**D**) are for a noninteractive model of community assembly assuming no numerical compensation by surviving species. Simulations (**E**), (**F**), (**G**), and (**H**) are for an interactive model that assumes full numerical compensation following extinction of competitors.

Our models predict that loss of species diversity leads to a decline in mean BMD, regardless of extinction scenario (Fig. 2). Note, however, that Fig. 2, A to H, depict strikingly different patterns, suggesting that changes in the BMD depend on extinction scenario. Indeed, the rate of change, the species richness at which the BMD first declines, the variance surrounding the relation (i.e., predictability of change), and the range of potential values all depend on how species go extinct (Table 1). These divergent patterns are best explained by examining the covariance between each species extinction risk and the biological traits that influence bioturbation (Fig. 3). To illustrate these patterns, we first focus on scenarios of extinction that involve no compensatory responses (i.e., the noninteractive model; Fig. 2, A, B, C, and D). Random extinction (Fig. 2A) produces a clear bifurcation of the BMD, with values determined by the presence (>4.0 cm) versus absence (<4.0 cm) of a single species—the burrowing brittlestar, *Amphiura filiformis*. The strong impact of *A. filiformis* on bioturbation is well documented (*27*). In this study, *A. filiformis* has a disproportionate impact (Fig. 3A) on bioturbation because it is consistently one of the most abundant species in Galway Bay (Fig. 3B) and has a high per capita effect that results from it being a large (Fig. 3C), highly mobile species. Consequently, changes in the BMD following extinction largely depend on whether *A. filiformis* is among the survivors.

When extinctions are ordered by species sensitivity to stress (Fig. 2B), estimated as the relative change in the abundance of species along a gradient of disturbance (*25*), the risk of extinction among species varies by a

factor of 215; yet, stress sensitivity for *A. filiformis* (−0.99, Fig. 3D) is near the median value for the community as a whole (−0.98), which explains why changes in the BMD are comparable to the scenario of random extinction (compare Fig. 2, A and B). This conclusion is confirmed by statistical comparisons of the mean and range of values (minimum and maximum) of the BMD, which show an identical change with species loss for both scenarios; and a comparison of the variability in BMD, which reveals only a marginal difference between scenarios (α = 0.0125; P = 0.01, Table 1).

For extinctions ordered by body size (Fig. 2C), probabilities of extinction were assumed to be proportional to mean species biomass to mimic the higher extinction risk generally faced by large-bodied organisms that often have small population sizes, have longer generation times, or are found at higher trophic levels (*17*, *28*). Body size varied by a factor of >500,000 among species and was positively correlated with per capita effects on bioturbation (r = 0.98, P < 0.01) but not abundance (r = −0.05, P = 0.56, even excluding *A. filiformis*, r = −0.08, P = 0.33). In this scenario, larger species (high per capita effects) tended to be lost before smaller species (low per capita effects), leading to a faster decline in the mean BMD compared with random extinction (compare Fig. 2, A and C; Table 1). The range of values of the BMD (minimum and maximum) and total variation (CV) also changed with species richness more quickly than for random extinctions (Table 1). This was not due to the loss of entire functional guilds composed of large species because there was considerable overlap in species body size, and thus extinction risk, among functional guilds (*25*). Rather, patterns were generally a consequence of the early extinction of *A. filiformis*, the 19th largest species, which produced a step change in the BMD at a species richness of ≈100.

Extinction risk is typically high for rare species, defined here as those with low local abundances, because small populations are more vulnerable to environmental and demographic stochasticity (*17*, *28*). They also often have narrow geographic ranges and/or high specialization, further compounding extinction risk (*28*–*30*). When we assumed extinction probability was inversely proportional to species density, rare species were >6000 times more likely to be lost than the most common species. Yet, because small populations typically contribute little to bioturbation (Fig. 3B), extinctions of rare species had little impact on the BMD, and ecosystem functioning was maintained until the loss of more abundant species, such as *A. filiformis* (lower bifurcation, Fig. 2D). Hence, some scenarios of extinction do not lead to appreciable loss of ecological function until a large proportion of species are lost.

Many studies suggest that when species go extinct from communities characterized by strong interactions, increases in the population size of species released from competition can compensate for loss of ecological function (*20*, *31*, *32*). Our models suggest that this is only true when the risk of extinction is not correlated with species functional traits. This is evident because compensatory responses only changed the probabilistic distribution of the BMD when species were lost at random (Fig. 2E) or in order of their sensitivity to stress (compare Fig. 2, A and E, and Fig. 2, B and F) (Table 1). However, when a species' risk of extinction covaried with its body size or abundance, compensatory responses did not alter the consequences of diversity loss (compare Fig. 2, C and G and Fig. 2, D and H) (Table 1). This is because when loss is ordered by body size, small species have little impact on bioturbation and cannot offset functions performed by larger species. When species are lost in order of rarity, even full compensation has no notable effect on the BMD because the proportional change in bioturbation is small. Thus, compensatory responses of surviving species do not necessarily stabilize ecological processes when the traits required for maintaining function simultaneously increase extinction risk.

We have used numerical models parameterized by data from a marine benthic community to show that species extinction is generally expected to reduce the depth of bioturbated sediments. Such changes might be expected to alter the fluxes of energy and matter that are vital to the global persistence of marine communities (*23*), a conclusion that corresponds to evolutionary patterns in the fossil record showing a close association between the frequency of anoxia and the diversification of marine soft-bottom communities (*33*). We have also shown that crucial details (mean, range, and predictabil-

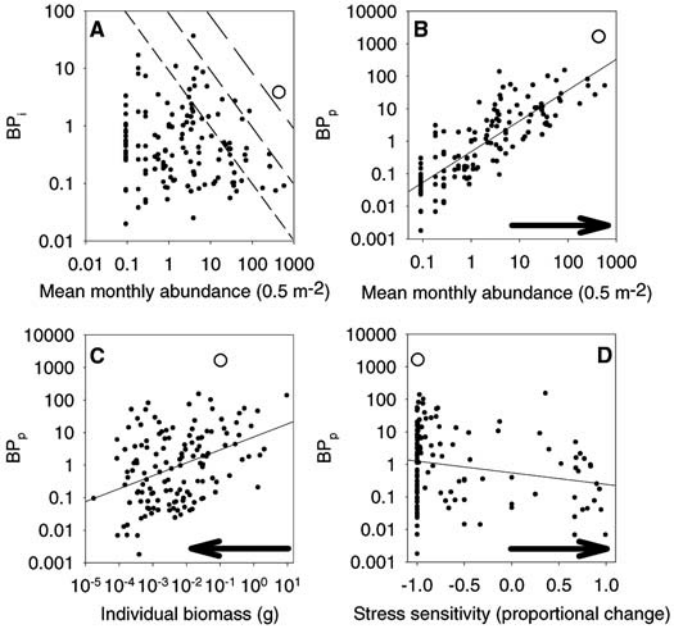


Fig. 3. The relation between per capita bioturbation, BP_i, and mean species abundance (**A**) reveals that at the population level (diagonal dashed lines, each an order of magnitude difference in bioturbation), most species contribute little to bioturbation (left of short-dashed line). Bioturbation is disproportionately affected by one large and highly active species, *Amphiura filiformis* (brittlestar, open circle). Population level bioturbation, BP_p, is proportional to species abundance (**B**) (r = 0.83, P < 0.001), body size (**C**) (r = 0.39, P < 0.001), and sensitivity to stress (**D**) (r = −0.2, P < 0.05). Arrows indicate order of extinctions.

Table 1. Comparisons of how bioturbation changes with species loss for each extinction scenario (stress, size, rarity) relative to a random model of extinction, and between the interactive and noninteractive models of community assembly. The asterisk (*) denotes significant differences, P < 0.0125 [set conservatively to correct for the number of comparisons (*25*)]. CV, coefficient of variation.

	Mean	CV	Minimum	Maximum
Comparison of random extinction to extinctions ordered by…				
Sensitivity to stress	χ^2_4 = 0.73	$F_{4, 1094}$ = 3.38*	χ^2_4 = 1.63	χ^2_4 = 0.23
Body size	χ^2_4 = 53.8*	$F_{4, 1094}$ = 42.8*	χ^2_4 = 15.1*	χ^2_4 = 15.1*
Rarity	χ^2_4 = 28.2*	$F_{4, 1094}$ = 250*	χ^2_4 = 97.6*	χ^2_4 = 3.8
Comparison of interactive to noninteractive model for extinctions that are…				
Random	χ^2_2 = 35.07*	$F_{2, 274}$ = 629*	χ^2_2 = 30.94*	χ^2_2 = 10.37*
Ordered by sensitivity to stress	χ^2_2 = 25.76*	$F_{2, 274}$ = 307*	χ^2_2 = 20.94*	χ^2_2 = 10.19*
Ordered by body size	χ^2_2 = 7.42	$F_{2, 274}$ = 166*	χ^2_2 = 10.71*	χ^2_2 = 5.56
Ordered by rarity	χ^2_2 = 1.38	$F_{2, 274}$ = 13.9*	χ^2_2 = 0.69	χ^2_2 = 0.50

ity of change) of how bioturbation changes following extinction depend on the order in which species are lost, because extinction risk is frequently correlated with life-history traits that determine the intensity of bioturbation. This finding is important because it argues that the particular cause of extinction ultimately governs the ecosystem-level consequences of biodiversity loss. Therefore, if we are to predict the ecological impacts of extinction and if we hope to protect coastal environments from human activities that disrupt the ecological functions species perform, we will need to better understand why species are at risk and how this risk covaries with their functional traits.

References and Notes

1. G. C. B. Poore, G. D. F. Wilson, *Nature* **361**, 597 (1993).
2. P. G. Falkowski *et al.*, *Science* **281**, 200 (1998).
3. P. Vitousek, H. Mooney, J. Lubchenco, J. Melillo, *Science* **277**, 494 (1997).
4. R. E. Turner, N. N. Rabalais, *Nature* **368**, 619 (1994).
5. J. B. C. Jackson *et al.*, *Science* **293**, 629 (2001).
6. M. Jenkins, *Science* **302**, 1175 (2003).
7. D. Malakoff, *Science* **277**, 486 (1997).
8. O. E. Sala *et al.*, *Science* **287**, 1170 (2000).
9. M. C. Emmerson, M. Solan, C. Emes, D. M. Paterson, D. Raffaelli, *Nature* **411**, 73 (2001).
10. C. L. Biles *et al.*, *J. Exp. Mar. Biol. Ecol.* **285**, 165 (2003).
11. S. G. Bolam, T. F. Fernandes, M. Huxham, *Ecol. Monogr.* **72**, 599 (2002).
12. D. Raffaelli, M. Emmerson, M. Solan, C. Biles, D. Paterson, *J. Sea Res.* **49**, 133 (2003).
13. B. Schmid *et al.*, in *Biodiversity and Ecosystem Functioning*, M. Loreau, S. Naeem, P. Inchausti, Eds. (Oxford Univ. Press, Oxford, 2002), pp. 61–75.
14. D. S. Srivastava, *Oikos* **98**, 351 (2002).
15. C. R. Tracy, T. L. George, *Am. Nat.* **139**, 102 (1992).
16. S. L. Pimm, H. L. Jones, J. Diamond, *Am. Nat.* **132**, 757 (1988).
17. M. L. McKinney, *Annu. Rev. Ecol. Syst.* **28**, 495 (1997).
18. D. Pauly, V. Christensen, J. Dalsgaard, R. Froese, F. Torres Jr., *Science* **279**, 860 (1998).
19. J. E. Duffy, *Ecol. Lett.* **6**, 680 (2003).
20. A. R. Ives, B. J. Cardinale, *Nature* **429**, 174 (2004).
21. M. D. Smith, A. K. Knapp, *Ecol. Lett.* **6**, 509 (2003).
22. M. Jonsson, O. Dangles, B. Malmqvist, F. Guérold, *Proc. R. Soc. London B Biol. Sci.* **269**, 1047 (2002).
23. U. Witte *et al.*, *Nature* **424**, 763 (2003).
24. K. S. Johnson *et al.*, *Nature* **398**, 697 (1999).
25. Materials and methods are available as supporting material on *Science* Online.
26. J. L. Ruesink, D. S. Srivastava, *Oikos* **93**, 221 (2001).
27. M. Solan, R. Kennedy, *Mar. Ecol. Prog. Ser.* **228**, 179 (2002).
28. J. H. Lawton, in *Population Dynamic Principles*, J. H. Lawton, R. M. May, Eds. (Oxford Univ. Press, Oxford, 1995), pp. 147–163.
29. K. F. Davies, C. F. Margules, J. F. Lawrence, *Ecology* **85**, 265 (2004).
30. C. N. Johnson, *Nature* **394**, 272 (1998).
31. D. D. Doak *et al.*, *Am. Nat.* **151**, 264 (1998).
32. J. M. Fischer, T. M. Frost, A. R. Ives, *Ecol. Appl.* **11**, 1060 (2001).
33. D. K. Jacobs, D. R. Lindberg, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9396 (1998).

34. We thank J. E. Duffy, J. D. Fridley, A. Hector, A. R. Ives, S. Naeem, O. L. Petchey, K. J. Tilmon, D. A. Wardle, and J. P. Wright for comments and the BIOMERGE Second Adaptive Synthesis Workshop for insightful discussion. Supported by BIOMERGE (Biotic Mechanisms of Ecosystem Regulation in the Global Environment)—an NSF-funded research coordination network (to S. Naeem).

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1177/DC1
Materials and Methods
Equations S1 and S2

23 July 2004; accepted 23 September 2004

Bushmeat Hunting, Wildlife Declines, and Fish Supply in West Africa

Justin S. Brashares,[1,2]* Peter Arcese,[3] Moses K. Sam,[4] Peter B. Coppolillo,[5] A. R. E. Sinclair,[6] Andrew Balmford[1,7]

The multibillion-dollar trade in bushmeat is among the most immediate threats to the persistence of tropical vertebrates, but our understanding of its underlying drivers and effects on human welfare is limited by a lack of empirical data. We used 30 years of data from Ghana to link mammal declines to the bushmeat trade and to spatial and temporal changes in the availability of fish. We show that years of poor fish supply coincided with increased hunting in nature reserves and sharp declines in biomass of 41 wildlife species. Local market data provide evidence of a direct link between fish supply and subsequent bushmeat demand in villages and show bushmeat's role as a dietary staple in the region. Our results emphasize the urgent need to develop cheap protein alternatives to bushmeat and to improve fisheries management by foreign and domestic fleets to avert extinctions of tropical wildlife.

The trade in bushmeat for human consumption is a key contributor to local economies throughout the developing world (*1*, *2*), but it is also among the greatest threats to the persistence of tropical wildlife (*1*–*4*). Efforts to manage the bushmeat trade are built on the premise that bushmeat consumption is driven by protein limitation. Thus, it is assumed that increases in livestock and agricultural production will reduce human reliance on wild sources of food (*5*–*7*). Although it makes intuitive and economic sense that consumption of wild meat would be linked to the availability of alternative sources of protein, there is little empirical evidence to support this assumption, particularly at large geographic scales (*1*, *5*–*7*). Furthermore, contrary to predictions of the "protein limitation" hypothesis, unsustainable consumption of wildlife remains a problem even in many relatively prosperous countries (*1*). Identifying bushmeat's value as a dietary staple versus a nonessential good is vital for targeting conservation interventions and, equally important, for predicting the impacts of wildlife declines on human livelihoods.

We evaluated the protein limitation hypothesis by comparing annual rates of decline for 41 species of wild carnivores, primates, and herbivores (table S1) in six nature reserves in Ghana with supply of fish in the region from 1970 to 1998. As is the case across the tropics, wild terrestrial mammals are used as a secondary source of animal protein in Ghana, and they comprise the chief commodities in a regional bushmeat trade estimated conservatively at 400,000 tons per year (*8*). Marine and freshwater fish are the primary source of animal protein consumed in West Africa, and the fisheries sector directly and indirectly accounts for up to one quarter of the workforce in the region (*9*, *10*). From 1965 to 1998, the supply of harvested fish in Ghana (Fig. 1A) ranged from 230,000 to 480,000 tons annually and varied by as much as 24% between consecutive years (*11*). Here, we test a prediction of the protein limitation hypothesis that years with low fish supply will show larger declines in biomass of terrestrial mammals, suggesting a transfer of harvest pressure and consumption between these resources. We also test for evidence of a mechanism underpinning such a transfer by examining (i) rates of hunting in nature reserves, (ii) sales and price data from local markets, and (iii) spatial trends in correlations of fish supply and wildlife declines.

[1]Conservation Biology Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK. [2]Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA. [3]Centre for Applied Conservation Research, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [4]Ghana Wildlife Division, Accra, Ghana. [5]Wildlife Conservation Society, Bronx, NY 10460, USA. [6]Centre for Biodiversity Research, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [7]Percy Fitz Patrick Institute of African Ornithology, University of Cape Town, Rondebosch 7701, Cape Town, South Africa.

*To whom correspondence should be addressed. E-mail: brashares@nature.berkeley.edu

In support of the prediction that annual standing biomass of large mammals would be linked positively with the annual supply of marine and freshwater fish, we found that changes in annual biomass of terrestrial mammals from 1970 to 1998 were closely related to annual per capita fish supply. Years with a lower-than-average supply of fish had higher-than-average declines in mammal biomass, and vice versa (Fig. 1B) (*12*). In contrast, fish supply and wildlife declines were unrelated to other potential explanatory factors, including annual rainfall, land and water temperatures, political cycles, oil prices, and gross national product ($P \geq 0.19$ for each term in multiple-regression models) (*13*). This correlative support for the protein limitation hypothesis is further supported by three additional analyses.

First, our working hypothesis suggests that the observed link between fish supply and wildlife decline occurs because bushmeat hunting and consumption increased when fish became scarce. In support of this suggestion, we found that annual counts of hunters observed by wildlife rangers in five nature reserves in Ghana (*13*) were related negatively to per capita fish supply from 1976 to 1992 (Fig. 2A). Annual counts of hunters were also closely related to annual rates of wildlife decline in these same nature reserves ($R = 0.76$, $n = 17$, $P < 0.01$). Thus, hunters were more common in reserves in years when fish supply was low, and these increases in hunters were linked to accelerated declines of wildlife.

Second, if annual variation in fish supply and bushmeat hunting are linked causally, we would expect that the availability of bushmeat in local markets would be related negatively to the supply of fish (*5*). In support of this prediction, we found that monthly supply of fish in 12 local markets in northern, central, and eastern Ghana from 1999 to 2003 (*13*) was related negatively to the volume of bushmeat sold in these markets (Fig. 2B). In addition, the price of fish sold in markets was closely and negatively related to local fish supply ($R = 0.73$, $n = 52$, $P < 0.01$) and positively related to the volume of bushmeat sold ($R = 0.48$, $n = 52$, $P < 0.01$). The strong negative correlation between fish price and quantity sold, combined with the positive correlation between fish price and bushmeat sales, is consistent with the idea that variation in fish supply drove bushmeat sales. Comparing monthly fish price in markets with the bushmeat sales in the following month yielded even stronger correlations, again suggesting that bushmeat sales were driven by fish availability and price more so than the reverse case (fig. S1). These results show a substitution of wildlife for fish at the local scale. Taken together with the observation of increased bushmeat hunting during periods of fish scarcity, these results also support our suggestion of a causal, macroscale link between fish supply and wildlife declines (Fig. 1).

Third, more than half of Ghana's human population of 20 million resides within 100 km of the coast, where the majority of employment and dietary protein are derived from fishing (*10*). Poor fish harvests result in reduced income and food for coastal communities and reduce the availability of fish throughout the region (*9*, *14*). The widespread loss of jobs and income associated with poor fish harvests also may lead some portion of households to rely on bushmeat hunting both for income and sustenance. If fish supply and bushmeat consumption are linked causally, it follows that the transfer of harvest pressure between aquatic and terrestrial resources would be most evident in

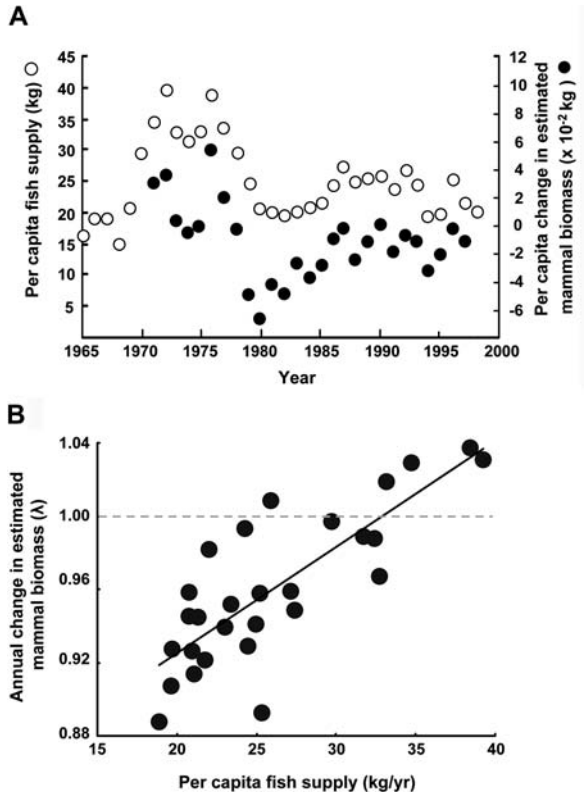


Fig. 1. Year-to-year change in estimated biomass of 41 large mammal species was linked closely to annual harvest of marine and freshwater fish in Ghana ($R = 0.73$, $n = 28$ years, $P < 0.001$). (**A**) Time series plots of annual fish supply and change in estimated mammal biomass. (**B**) Conventional plot of data shown in (A). The trend line describes the equation $y = 0.0058x + 0.81$. Values of annual fish supply [from (*11*)] represent landings plus imports and minus exports. Biomass of large mammals was calculated for each year by multiplying the number of animals observed in ~700 walking counts of 10 to 15 km each (*17*) by species-specific body weights. The products of these calculations were then summed across all species.

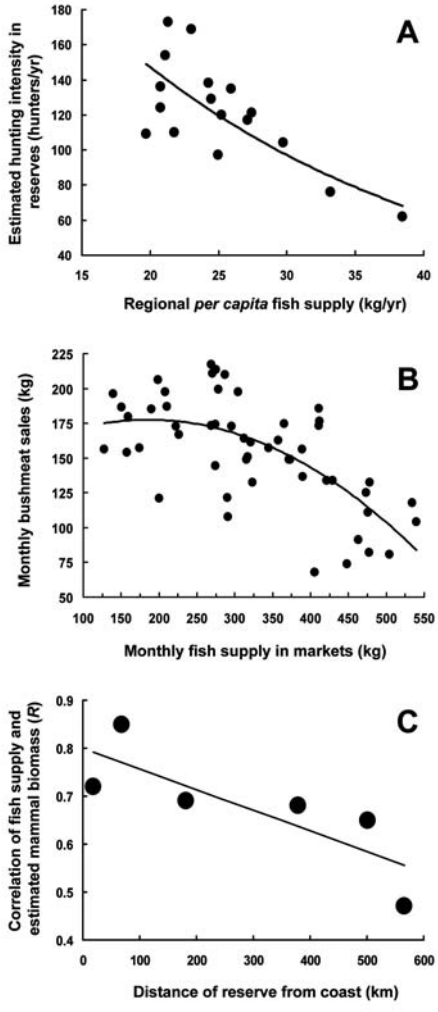


Fig. 2. Links between fish supply and bushmeat hunting and consumption are evident in observations that (**A**) annual counts of hunters in five terrestrial reserves in Ghana from 1976 to 1992 were related negatively to supply of fish in the region ($R = -0.52$, $n = 17$, $P = 0.03$); (**B**) monthly sales of bushmeat in 12 rural markets in Ghana were related negatively to local fish supply ($R = -0.61$, $n = 52$, $P < 0.01$); and (**C**) fish supply and wildlife declines were related most closely in reserves occurring nearest to the coast ($R = 0.81$, $n = 6$, $P = 0.05$).

coastal areas where reliance on fish for both income and animal protein is greatest. We tested this last prediction by repeating the analysis in Fig. 1 separately for each of six nature reserves in Ghana. We found the strongest link between annual variation in marine and freshwater fish supply and annual change in mammal biomass in reserves near the coast and weaker, though still significant, linkages for reserves farther inland (Fig. 2C).

These three lines of evidence indicate that fish supply is linked negatively to the price of fish, the number of wildlife hunters, and the sales and supply of bushmeat in local markets. Our results also show that the substitution of fish for bushmeat occurs most intensively close to the coast, where fish are more important as sources of food and income. All of these findings are consistent with the protein limitation hypothesis and inconsistent with the notion that bushmeat in Ghana is primarily a nonessential good (summarized in fig. S2).

Our results provide clear evidence to suggest that the outcomes of programs aimed at promoting economic development, food security, and the conservation of biological diversity in Ghana, and perhaps elsewhere in Africa, will be closely linked. First, the close correlation between hunting pressure, markets, and long-term trends in wildlife abundance suggests strongly that the persistence of the more than 400 species of terrestrial vertebrates that supply the bushmeat trade in West Africa will depend ultimately on the availability of affordable alternative protein sources for the region's growing human population. Second, our failure to conserve existing wildlife populations as core sources for managed, sustainable harvests could have serious deleterious effects on the stability of the long-term human food supply and the livelihoods of bushmeat hunters and sellers. Our findings and those of others suggest that the harvest of terrestrial wildlife can buffer the impact of environmental or other shocks by providing animal protein and income in times of economic hardship or food scarcity (*2*, *15*, *16*). However, marked declines in large mammal abundance and marine and freshwater fish stocks documented in the region over the past 30 years now suggest that this buffer system can no longer be sustained (*14*, *17–20*).

From 1970 to 1998, the biomass of 41 species of mammals in nature reserves in Ghana declined by 76% (Fig. 3), and 16 to 45% of these species became locally extinct (*17*). Similarly, trawl surveys conducted in the Gulf of Guinea since 1977 and other regional stock assessments estimate that fish biomass in nearshore and offshore waters has declined by at least 50% (Fig. 3). At the same time, a threefold increase in human populations in the region since 1970 has resulted in per capita declines in fish supply, despite steady increases in regional fish harvests (*11*, *14*). These sharp declines in terrestrial wildlife and marine fish suggest that stocks in this region may face imminent collapse (*9*, *18*). The consequences of collapse of either fish or terrestrial wildlife are daunting and may be felt immediately as widespread human poverty and food insecurity in the region (*14*). Reduced fish stocks have already severely damaged the region's artisanal fisheries sector (*14*, *21*), and recent collapses of mammal populations in some areas of West Africa have been linked to geographic patterns of poverty and malnourishment (*8*, *17*). Agricultural production is a third potentially critical, though poorly understood, factor linking human food supply to biodiversity conservation in the region (*16*).

One management response to the potential collapse of fish and terrestrial wildlife stocks in West Africa is to build up regional livestock and agriculture sufficiently to alleviate pressure on overexploited wild resources (*7*). However, such efforts could take decades to implement and face enormous economic, regulatory, and political hurdles. Thus, more immediate plans to enhance the sustainability of wild protein sources are required. One immediate route to increasing production and sustainability of domestic fisheries, and thereby reducing pressure on terrestrial wildlife, is to limit the access of large and heavily subsidized foreign fleets to fish off West Africa (*18–24*). Declines of fish stocks in nearshore and offshore waters of West Africa have coincided with more than 10-fold increases in regional fish harvests by foreign and domestic fleets since 1950 (*11*). The European Union (EU) has consistently had the largest foreign presence off West Africa, with EU fish harvests there increasing by a factor of 20 from 1950 to 2001 (fig. S3). Furthermore, EU financial support of its foreign fleet increased from about $6 million in 1981 to more than $350 million in 2001 (fig. S3), with the effect of artificially increasing the profitability of fishing in African waters for EU boats, despite declining fish stocks (*22*). West African commercial fleets also have expanded considerably since 1950 (fig. S3) and there is no guarantee that reductions of foreign catches would not be taken up by increased domestic fishing. However, even short-term increases in the domestic supply of fish both for commercial export and local consumption may enhance regional economies (*14*) and ease exploitation of terrestrial wildlife resources. Over the longer term, intensive management to enhance fish stocks and stabilize harvests must become a regional conservation and economic priority.

A second route to increase the sustainability of fish and wildlife harvests could come by enhancing the protection of harvested marine and terrestrial resources. Pirate fishing vessels from foreign ports are abundant in West African waters and illegally extract fish of the highest commercial value while, like many commercial fleets, dumping 70 to 90% of their haul as by-catch (*9*, *18*). Increased policing of exclusive fishing zones and enforcement of existing quotas and tariffs for commercial fleets should reduce exploitation and provide an immediate boost to marine resources available to local fisheries (*14*, *19*). On land, wildlife has persisted at near historic levels in inaccessible and well-protected areas of West Africa's nature reserves (*4*, *17*). Increasing the size, number, and protection of wildlife reserves in the region may not offer a long-term solution to concerns over human livelihoods and protein supply, but it is likely to offer the most immediate prospects for slowing the region's catastrophic wildlife decline.

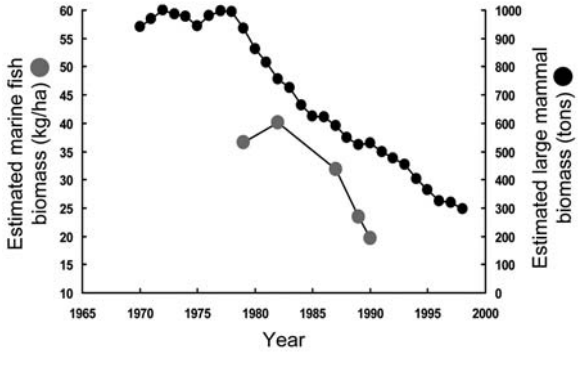


Fig. 3. Estimates of marine fish biomass in the Gulf of Guinea (gray circles) and large mammal biomass in Ghana (black circles). Estimates of fish biomass are from trawl surveys (*24, 25*). Analyses of fisheries catch data with ecosystem models indicate that fish biomass in coastal West and Northwest Africa has declined by a factor of 13 since 1960 (*20*). Estimates of mammal biomass are based on abundances of 41 species observed in ~700 wildlife counts per year in six nature reserves (*17*) (see map, fig. S4).

References and Notes

1. J. G. Robinson, E. L. Bennett, Hunting for Sustainability in Tropical Forests (Columbia Univ. Press, New York, 2000).
2. E. J. Milner-Gulland *et al.*, *Trends Ecol. Evol.* **18**, 351 (2003).
3. J. G. Robinson, K. H. Redford, E. L. Bennett, *Science* **284**, 595 (1999).
4. World Conservation Union, International Union for Conservation of Nature and Natural Resources (IUCN) Red List of Threatened Animals (IUCN, Gland, Switzerland, 2000).
5. D. S. Wilkie, R. A. Godoy, *Science* **287**, 975 (2000).
6. E. L. Bennett, *Conserv. Biol.* **16**, 588 (2002).

7. Organisation for Economic Co-operation and Development (OECD), *Shaping the 21st Century: The Contribution of Development Cooperation* (OECD, Paris, 1996).
8. Y. Ntiamoa-Baidu, *Wildlife Development Plan: 1998–2003* (Wildlife Department, Accra, Ghana, 1998).
9. United Nations Environment Programme (UNEP), *Africa Environment Outlook*; available at www.unep.org/aeo.
10. Food and Agriculture Organization (FAO), *Country Profiles: Ghana*; available at www.fao.org/countryprofiles/index.asp?iso3=GHA.
11. Food and Agriculture Organization (FAO), *Fisheries Databases*; available at www.fao.org/fi/statist/statist.asp.
12. Statistics are based on a linear regression of annual change in mammal biomass [calculated as $(kg_{t+1})/kg_t$] against per capita fish harvest. Regressing per capita change in mammal biomass [i.e., $(kg_{t+1} - kg_t)$/human population $_{t+1}$] against per capita fish catch gave a similar result (adjusted $R^2 = 0.52$, $P < 0.001$).
13. Materials and methods are available as supporting material on *Science* Online.
14. J. Atta-Mills, J. Alder, U. R. Sumaila, *Nat. Resour. Forum* **28**, 13 (2004).
15. C. B. Barrett, P. Arcese, *Land Econ.* **74**, 449 (1998).
16. G. J. S. Dei, *Ecol. Food Nutr.* **22**, 225 (1989).
17. J. S. Brashares, P. Arcese, M. K. Sam, *Proc. R. Soc. Lond. Ser. B.* **268**, 2473 (2001).
18. Food and Agriculture Organization (FAO), *The State of World Fisheries and Aquaculture 2002*; available at www.fao.org/docrep/005/y7300e/y7300e00.htm.
19. D. Pauly *et al.*, *Nature* **418**, 689 (2002).
20. V. Christensen *et al.*, in *Pêcheries Maritimes, Ecosystèmes et Sociétés: Un Demi-Siècle de Changement*, B. A. Moctar, P. Chavance, D. Gascuel, M. Vakily, D. Pauly, Eds. (Institut de Recherche pour le Developpement, Paris, 2004).
21. World Wide Fund for Nature (WWF), *West Africa Puts EU To Shame* (WWF European Policy Office, Brussels, 2001).
22. V. M. Kaczynski, D. L. Fluharty, *Mar. Policy* **26**, 75 (2002).
23. S. L. Pimm *et al.*, *Science* **293**, 2207 (2001).
24. Sea Around Us Project (SAUP), *Web Products: Marine Database*; available at www.seaaroundus.org.
25. Ghan Marine Fisheries Research Division (MFRD), Oceangraphic Data Center, *Marine database*; avilable at www.ioc.unesco.org/odinafrica/contnets.php?=217.
26. We thank the Ghana Wildlife Division for permission to work in reserves and access to data, and we thank J. Atta-Mills for discussion. C. Kresge, P. Kresge, J. Mason, B. Volta, N. Ankudey, D. Boateng, L. Lanto, and G. Agbango provided assistance in Ghana, and C. Barrett, J. Hellmann, D. Pauly, I. Watson, J. Smith, V. Christensen, E. J. Milner-Gulland and four anonymous reviewers gave many helpful suggestions. Supported by NSF INT-0301935 (J.S.B.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/306/5699/1180/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References

7 July 2004; accepted 7 October 2004

The Genetic Basis of Singlet Oxygen–Induced Stress Responses of *Arabidopsis thaliana*

Daniela Wagner,[1]*† Dominika Przybyla,[1]* Roel op den Camp,[1] Chanhong Kim,[1] Frank Landgraf,[1] Keun Pyo Lee,[1] Marco Würsch,[1] Christophe Laloi,[1] Mena Nater,[1] Eva Hideg,[2] Klaus Apel[1]‡

Plants under oxidative stress suffer from damages that have been interpreted as unavoidable consequences of injuries inflicted upon plants by toxic levels of reactive oxygen species (ROS). However, this paradigm needs to be modified. Inactivation of a single gene, *EXECUTER1*, is sufficient to abrogate stress responses of *Arabidopsis thaliana* caused by the release of singlet oxygen: External conditions under which these stress responses are observed and the amounts of ROS that accumulate in plants exposed to these environmental conditions do not directly cause damages. Instead, seedling lethality and growth inhibition of mature plants result from genetic programs that are activated after the release of singlet oxygen has been perceived by the plant.

Abiotic stress conditions limit the ability of plants to use light energy for photosynthesis, often reducing their growth and productivity and causing photooxidative damages (*1*–*3*). The emergence of these stress symptoms has been closely associated with the enhanced production of several ROS (*4*, *5*). Because different ROS are generated simultaneously, it is difficult to determine the biological activity and mode of action for each of these ROS separately. In order to address this problem, one would need to find conditions under which only one specific ROS is generated at a given time, within a well-defined subcellular compartment, and which also triggers a visible stress response that is easy to score.

Recently, we have isolated the conditional *flu* mutant of *Arabidopsis thaliana* that fulfills these requirements (*6*). The mutant generates singlet oxygen in plastids in a controlled and noninvasive manner. Immediately after the release of singlet oxygen, mature *flu* plants stop growing, whereas seedlings bleach and die (*6*). Here, we demonstrate that the two stress responses, growth inhibition and seedling lethality, do not result from physicochemical damage caused by singlet oxygen during oxidative stress but are caused by the activation of a genetically determined stress response program.

We set out to identify such a genetic program by identifying second-site mutations that abrogate either one or both of the two stress responses of the *flu* mutant. Three different groups of second-site mutations could be distinguished (*7*) (fig. S1A). One of these groups contained 15 mutants that behaved like wild type when kept under nonpermissive light-dark conditions (*7*) [group III (fig. S1, B to D)]. Allelism tests and mapping revealed that they were allelic, representing a single locus that was named *EXECUTER1*. In contrast to wild-type plants but like *flu*, the *executer1/flu* double mutant accumulated free protochlorophyllide (Pchlide) in the dark (Fig. 1, A to C, and fig. S1B). After transfer to the light, *executer1/flu* generated singlet oxygen in amounts similar to those of *flu* (Fig. 1, F to H) but grew like wild type when kept under nonpermissive light-dark cycles (Fig. 1, A to C). The second stress reaction of *flu* to the release of singlet oxygen is an inhibition of growth. In *flu* plants, the growth rate was reduced immediately after the beginning of reillumination (Fig. 1D). The *executer1/flu* plants, however, grew like wild-type plants (Fig. 1D). Growth inhibition of *flu* plants was particularly striking when plants were transferred to repeated light-dark cycles, whereas *executer1/flu* continued to grow like wild-type plants (Fig. 1E). All three plant lines grew equally well under continuous light (fig. S2).

As a first step toward the functional characterization of EXECUTER1, we used a map-based cloning strategy to isolate the *EXECUTER1* gene. *EXECUTER1* was genetically mapped on chromosome IV on a genomic fragment of about 90 kb (Fig. 2A). A contig consisting of 11 cosmid clones that encompassed this chromosomal region was generated (Fig. 2A), and the ability to complement the *executer1* mutation was tested (*7*). Seedlings of the double mutant transformed with the genomic DNA of the cosmid clone 44 that contained a wild-type copy of *EXECUTER1* died like *flu* seedlings when grown under nonpermissive dark-light conditions, whereas seedlings of plants transformed with genomic DNA of other cosmid clones grew like seedlings of the original *executer1/flu* parental line (Fig. 2B).

The second test was done with mature T2 plants transformed with DNA of cosmid

[1]Institute of Plant Sciences, Plant Genetics, Swiss Federal Institute of Technology (ETH), CH-8092 Zurich, Switzerland. [2]Institute of Plant Biology, Biological Research Center, Hungarian Academy of Sciences, H-6701 Szeged, Hungary.

*These authors contributed equally to this work.
†Deceased 24 February 2004.
‡To whom correspondence should be addressed. E-mail: klaus.apel@ipw.biol.ethz.ch

clone 44. The original *executer1/flu* mutant plants were not visibly affected by a shift to light-dark cycles and continued to grow like wild-type plants, but *executer1/flu* plants complemented with the cosmid 44 DNA were indistinguishable from the *flu* mutant in that they ceased growth and their leaves developed necrotic lesions (Fig. 2C). Both assays demonstrate that the rapid bleaching of *flu* seedlings and the inhibition of growth after the release of singlet oxygen are not because of the toxicity of this ROS and do not reflect photooxidative damage and injury, but instead result from the activation of genetically controlled responses that require the activity of the *EXECUTER1* gene.

The complementing genomic fragment of cosmid clone 44 has a size of about 14,000 base pairs (bp) and contains three open reading frames (ORFs). One of them, At4g33630, was identified as *EXECUTER1* (*7*) (fig. S3A). The ORF of the *EXECUTER1* cDNA predicts a protein of 684 amino acids with a molecular mass of 76,534.9 daltons. It is unrelated to known proteins, but its N-terminal part resembles import signal sequences of nuclear-encoded plastid proteins (fig. S3B). This prediction was confirmed experimentally (fig. S4, A to C). In all higher plants for which expressed sequence tag sequence data were available, including major crop plants, *EXECUTER1* homologs could be found (fig. S3C). Thus, EXECUTER1 represents a highly conserved plastid protein that seems to enable higher plants to perceive singlet oxygen as a stress signal that activates a genetically determined stress response program.

The physiological role of the *EXECUTER1* gene in wild-type plants was assessed by first isolating *executer1* mutant plants that no longer carried copies of the mutated *flu* gene (*7*). Wild-type and *executer1* plants were then exposed to higher light intensities in the presence of 3-(3, 4-dichlorphenyl)-1,1-dimethylurea (DCMU), which is known to stimulate the release of singlet oxygen (*8*). The effect of the *executer1* mutation on singlet oxygen–mediated cell death was assessed by floating cut leaves on solutions with increasing concentrations of DCMU, ranging from 25 nM to 25 μM, first in the dark for 30 min and then under light (950 μmol photons m^{-2} s^{-1}) for 24 hours. Inhibition of photosystem II (PSII) by DCMU was similar in wild types and *executer1* mutants, as indicated by the concurrent decline of the maximum efficiency of PSII at increasing DCMU concentrations in both leaf samples (Fig. 3A).

Cell death and membrane damage in cut leaves were estimated by measuring ion leakage. When leaves of wild type and *executer1* were tested in the absence of DCMU, ion leakage reached about 10% of the maximum after boiling the leaves (Fig. 3B). In the presence of 25 nM and 250 nM DCMU, ion leakage in leaves of wild type increased strongly to 40 and 55%, respectively, whereas in leaves of the *executer1* mutant ion leakage remained at a low level, similar to that in the water control (Fig. 3B). Only when the concentration of DCMU was further increased to 25 μM were leaves of both wild type and *executer1* almost completely damaged, with ion leakage being close to 100% after 24 hours of illumination (Fig. 3B). However, even under these harsher conditions, the increase of cell damage over time in DCMU-treated leaves of wild type occurred much faster than the damage in leaves of *executer1* (*9*).

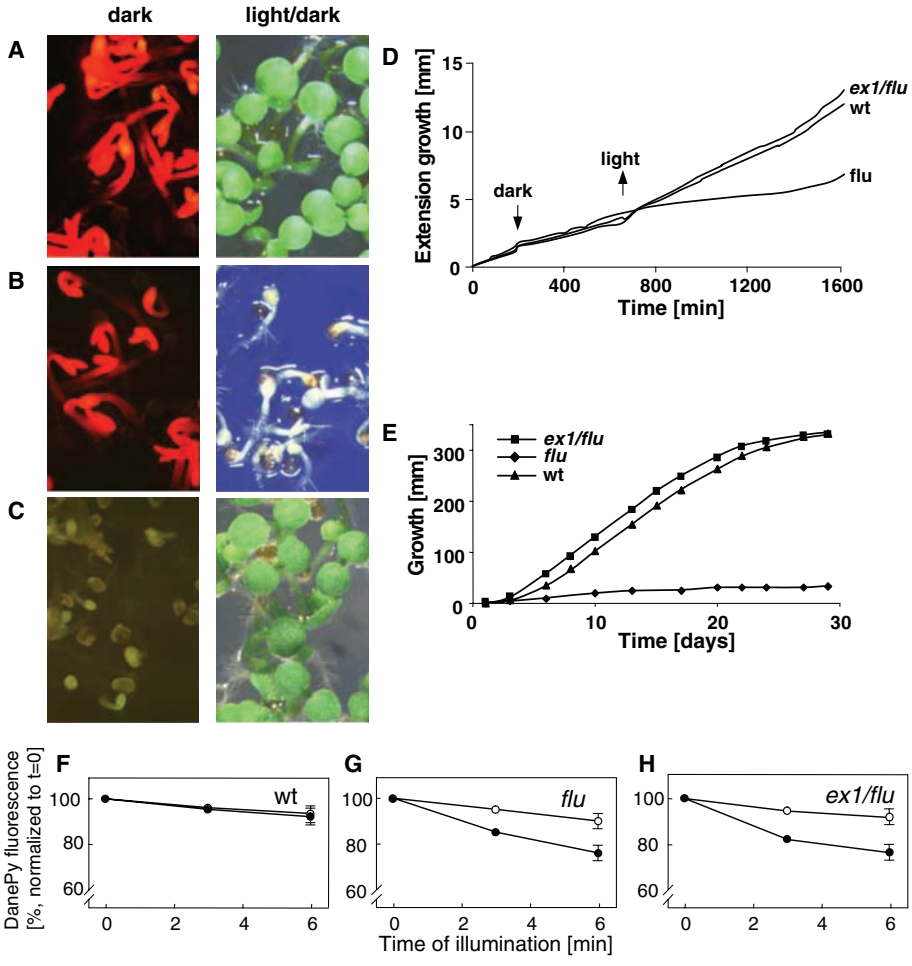


Fig. 1. A comparison of singlet oxygen production, cell death, and growth inhibition in *executer1/flu*, *flu*, and wild type (WT). Etiolated seedlings of *flu* (**B**) and *executer1/flu* (**A**) overaccumulated similar amounts of free Pchlide, as indicated by the bright red fluorescence, in contrast to etiolated WT seedlings (**C**). Cell death of *flu* seedlings (B) and growth inhibition of mature *flu* plants (♦) grown under nonpermissive 16 hours light–8 hours dark conditions were blocked by the *executer1* mutation (■) (**D** and **E**). Results from WT control plants (▲) are also shown in (D) and (E). (F to H) Generation of singlet oxygen. WT (**F**), *flu* (**G**), and *executer1/flu* (**H**) were grown under continuous light until they were ready to bolt. At this stage, plants were either transferred to the dark for 8 hours or kept under light. Cut leaves were infiltrated with dansyl-2,2,5,5-tetramethyl-2,5-dehydro-1H-pyrrole and subsequently illuminated with white light (100 μmol photons m^{-2} s^{-1}). Singlet oxygen trapping was measured as relative quenching of DanePy fluorescence (*6*). In WT controls, no difference in singlet oxygen production between plants exposed to continuous light and a dark-light shift could be found (H), whereas in leaves of *flu* and *executer1/flu* generation of singlet oxygen was enhanced in plants that had been kept in the dark before reillumination (●) but not in plants exposed to continuous light (○). *ex1*, *executer1*.

In most of the second-site mutants of *flu* that have been identified during our suppressor screen, only one of the two major singlet oxygen–mediated stress reactions was abolished, i.e., either growth inhibition or seedling lethality (fig. S1). In the *executer1/flu* mutant line, however, both visible stress responses were abrogated. One could argue that a mutated form of the EXECUTER1 protein may act as a scavenger of singlet oxygen and in this way may eliminate detrimental toxic effects of this ROS. However, in such a case, singlet oxygen should no longer be detectable in the *executer1/flu* mutant line. Furthermore, suppression of stress responses in *flu* by the *executer1* mutation

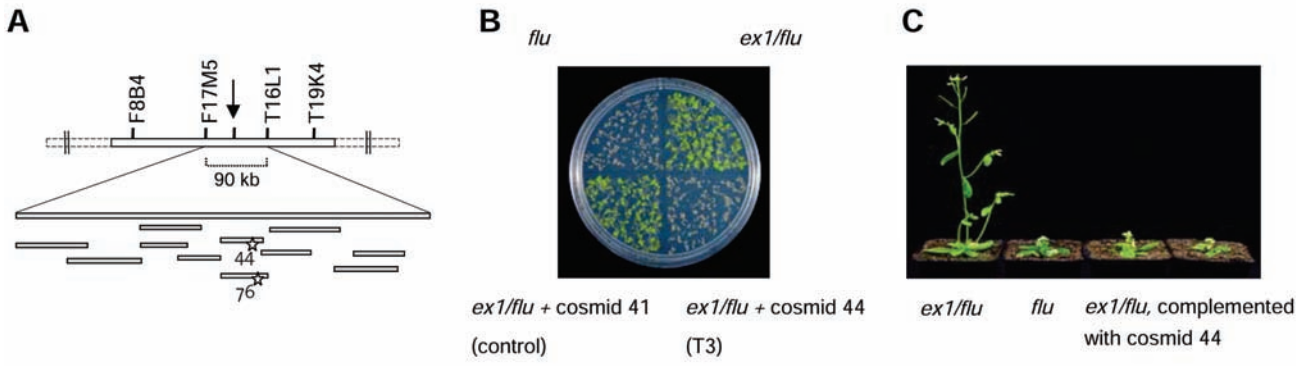


Fig. 2. Identification of the *EXECUTER1* gene. (**A**) Genetic and physical map of the DNA region on chromosome IV of *A. thaliana* that contains the *EXECUTER1* gene (arrow). The region between markers F17M5 and T16L1 is encompassed by 11 cosmid clones that were used to complement the *executer1-7/flu* double mutant. Cosmid clones 44 and 76, marked by stars, contained genomic DNA fragments that largely overlapped. The DNA fragment of cosmid 44 restored the cell death of seedlings (**B**) and the growth inhibition of mature plants of the parental *flu* line (**C**) when grown under nonpermissive light-dark cycles. Similar results were obtained with cosmid 76 (*15*).

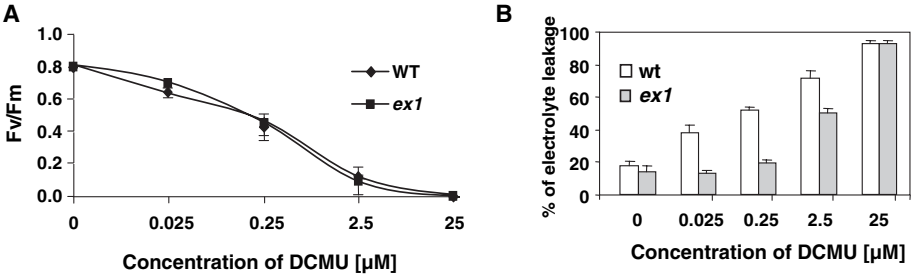


Fig. 3. Suppression of cell death in wild-type plants by the inactivation of *EXECUTER1*. (**A**) The effects of different concentrations of DCMU on the maximum efficiency of PSII in WT and *executer1-48* plants grown for 3 weeks under 16 hours light–8 hours dark conditions. (**B**) Selective suppression of cell death by *executer1* in DCMU-treated leaves kept at 950 μmol photons m^{-2} s^{-1}. The progression of cell death was determined by measuring the electrolyte leakage of cut leaves at different lengths of illumination. The ion leakage of leaves boiled for 25 min was taken as 100%. At concentrations of up to 0.25 μM DCMU, cell death was suppressed in *executer1*, whereas in WT plants it steadily increased with increasing DCMU concentrations. In leaves of the *flu* mutant kept in the dark for 8 hours before the experiment, the extent of ion leakage after 24 hours of reillumination at 100 μmol photons m^{-2} s^{-1} reached about 50% (*9*).

was not confined to allelic lines that synthesized modified EXECUTER1 proteins with single amino acid exchanges that could turn the mutated protein into a scavenger of singlet oxygen but occurred also in *executer1* allelic lines that were no longer able to synthesize this protein. Lastly, a mutation that confers an increased scavenging capacity to the *executer1/flu* mutant should be transmitted as a dominant trait, whereas all *executer1* alleles represent recessive mutations.

Inactivation of the EXECUTER1 protein did not only suppress the induction of death in *flu* seedlings grown under light-dark cycles but also in wild-type plants that were treated with DCMU. DCMU is known to stimulate the release of singlet oxygen in chloroplasts. It inhibits PSII and mimics photoinhibition by binding to the secondary quinone electron acceptor of PSII, Q_B, and inhibiting forward electron transport. Charge recombination in PSII favors the formation of a chlorophyll (Chl) triplet state that reacts with ground-state triplet oxygen to form $^1O_2$ (*10*).

Carotenoids of light-harvesting complexes effectively quench triplet Chl and singlet oxygen (*11*), but the β-carotenes bound to PSII reaction center fail to do so because they are localized too far away from the P680 Chl (*12*). Thus, continuous $^1O_2$ production seems to be an inherent property of PSII even under low-light conditions. The quenching of this singlet oxygen has been linked to the turnover of the D1 protein that is oxidized by singlet oxygen and apparently serves as a scavenger of this ROS (*13*, *14*). Excess amounts of singlet oxygen that cannot be quenched by the D1 protein and that interact with other targets within the vicinity of PSII may be the trigger that initiates singlet oxygen–mediated stress responses in wild-type plants (*14*). So far, these stress responses have been attributed to the toxicity of this ROS (*1*, *2*). However, as shown in our present work, the intensity and quality of these responses to light stress may range from necrotic reactions resulting from severe photooxidative damage to the activation of a genetically controlled cell death. At lower DCMU concentrations, ion leakage and membrane damage seem to result from the activation of the EXECUTER1-dependent cell-death program. At higher DCMU concentrations, this genetically controlled cell-death reaction is gradually masked by an EXECUTER1-independent cell-death reaction that seems primarily caused by the toxicity of elevated levels of singlet oxygen. In the past, it was not possible to distinguish between these two cell death reactions, because the genetically determined part remained unnoticed. With the identification of the *executer1* mutation, it is now possible to define conditions under which the genetically controlled cell death prevails.

References and Notes

1. J. Barber, B. Andersson, *Trends Biochem. Sci.* **17**, 61 (1992).
2. K. Apel, H. Hirt, *Annu. Rev. Plant Biol.* **55**, 373 (2004).
3. K. K. Niyogi, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 333 (1999).
4. M. J. Fryer, K. Oxborough, P. M. Mullineaux, N. R. Baker, *J. Exp. Bot.* **53**, 1249 (2002).
5. E. Hideg *et al.*, *Plant Cell Physiol.* **43**, 1154 (2002).
6. R. G. op den Camp *et al.*, *Plant Cell* **15**, 2320 (2003).
7. Materials and methods are available as supporting material on *Science* Online.
8. C. Fufezan, A. W. Rutherford, A. Krieger-Liszkay, *FEBS Lett.* **532**, 407 (2002).
9. D. Przybyla, unpublished results.
10. B. A. Diner, F. Rappoport, *Annu. Rev. Plant Biol.* **53**, 551 (2002).
11. R. Cogdell, H. A. Frank, *Photochem. Photobiol.* **63**, 257 (1996).
12. N. Kamiya, J.-R. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 98 (2003).
13. J. Sharma, M. Panico, J. Barber, H. R. Morris, *J. Biol. Chem.* **272**, 3935 (1997).
14. A. Trebst, *Z. Naturforsch.* **C58**, 609 (2003).
15. D. Wagner, unpublished results.
16. We are indebted to T. Fitzpatrick for critical reading, D. Rubli for art work, and M. Geier-Bächtold for editorial work. This study was supported by the Swiss Federal Institute of Technology and the Swiss NSF.

Supporting Online Material
www.sciencemag.org/cgi/content/full/306/5699/1183/DC1
Materials and Methods
Figs. S1 to S4

23 July 2004; accepted 21 September 2004

Microbial Factor-Mediated Development in a Host-Bacterial Mutualism

Tanya A. Koropatnick,[1] Jacquelyn T. Engle,[2] Michael A. Apicella,[3] Eric V. Stabb,[4] William E. Goldman,[2] Margaret J. McFall-Ngai[1,5]*

Tracheal cytotoxin (TCT), a fragment of the bacterial surface molecule peptidoglycan (PGN), is the factor responsible for the extensive tissue damage characteristic of whooping cough and gonorrhea infections. Here, we report that *Vibrio fischeri* also releases TCT, which acts in synergy with lipopolysaccharide (LPS) to trigger tissue development in its mutualistic symbiosis with the squid *Euprymna scolopes*. As components of PGN and LPS have commonly been linked with pathogenesis in animals, these findings demonstrate that host interpretation of these bacterial signal molecules is context dependent. Therefore, such differences in interpretation can lead to either inflammation and disease or to the establishment of a mutually beneficial animal-microbe association.

To date, molecules conserved among microbes, such as LPS and PGN, have been collectively described as "pathogen"-associated molecular patterns (PAMPs) (*1*). However, the majority of animal-microbe interactions are benign or mutualistic, raising the question: What role might such factors play in other types of host-microbe associations?

The reciprocal dialogue between partners in benign host-symbiont associations has been shown to be important for host tissue maturation (*2*–*4*), although the identification of the bacterial signals involved has proven elusive. The symbiosis between the Hawaiian bobtail squid *E. scolopes* and the luminous, Gram-negative bacterium *V. fischeri* offers the opportunity to decipher experimentally the precise dialogue between host and microbe partners. In this system, the bacterium colonizes epithelium-lined crypts within the host's light-emitting organ as a monospecific, extracellular symbiont (Fig. 1) (*5*).

Shortly after the juvenile squid emerges from the egg, the bacterial inoculum is gathered from the environment as seawater passes through the mantle cavity (*6*). Two prominent fields of ciliated epithelia on the surface of the squid's light organ facilitate the colonization process through ciliary motion and mucus shedding, resulting in the aggregation of symbiont cells above pores on the surface of the organ. Once aggregated, the symbionts migrate through the pores, down ducts, and into crypts that are located 100 to 200 μm from the surface epithelia (Fig. 1). Upon colonization of the crypts, some of the first processes triggered by the symbiont include the infiltration of macrophage-like hemocytes (blood cells) into the sinuses of the ciliated fields (Fig. 2A) and the induction of widespread apoptosis of the epithelial cells that compose these fields (*5*) (Fig. 2B). The most conspicuous response to light organ colonization is the extensive morphogenesis of the organ's surface, which culminates in the complete loss of the ciliated field 4 days after the initial colonization by the symbiont (Fig. 2C). In nature, only *V. fischeri* is capable of colonizing the crypts and signaling this morphogenesis, which serves to transform the organ from a morphology associated with the colonization process to one characteristic of the mature, functional organ. PGN in ambient seawater is known to trigger the shedding of mucus from the epithelial fields of the light organ (*7*), and LPS can induce low levels of early-stage apoptosis (*8*). However, the symbiont-derived factor(s) capable of triggering the full sequence of light organ morphogenesis have yet to be identified.

V. fischeri continuously sheds fragments of its surface in culture, and the crypt epithelial cells that interface closely with the symbiont cross-react with a monoclonal antibody to *V. fischeri* LPS (*9*). Cell surface fractions isolated from *V. fischeri* were sufficient to induce levels of hemocyte infiltration, apoptosis, and regression of the epithelial fields comparable to those induced by the intact symbiont (Fig. 2, D to F). Cell surface fractions from the nonsymbiotic, marine, Gram-negative bacterium *Pseudoalteromonas luteoviolacea* were also active (*9*).

V. fischeri PGN signaled levels of hemocyte infiltration comparable to those initiated by the intact symbiont, whereas *V. fischeri* LPS did not induce this cellular reaction (Fig. 2G). Alone, PGN did not trigger apoptosis, although LPS did stimulate it at low levels; however, together PGN and LPS acted synergistically to induce apoptosis at levels characteristic of those resulting from colonization by the intact symbiont (Fig. 2H). PGN, both alone and in synergy with LPS, also induced significant levels of epithelial regression (Fig. 2I). These findings are similar to those reported for certain pathogenic associations in which PGN and PGN fragments can induce macrophage activation (*10*) and infiltration (*11*) into inflamed host tissues. Likewise, purified PGN and LPS work in synergy to

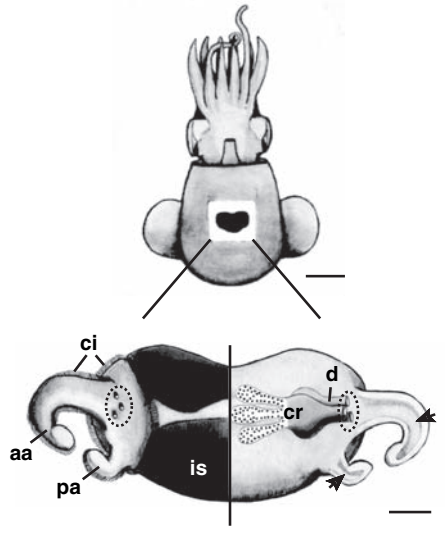


Fig. 1. *E. scolopes* possesses a light-emitting organ. (**Top**) A ventral view of a juvenile squid shows a window through the mantle to illustrate the position of the organ within the mantle cavity. Scale bar, 500 μm. (**Bottom**) An enlargement shows the details of the juvenile light organ morphology. The left half illustrates the surface with its ciliated epithelial field, which is composed of an anterior (aa) and posterior (pa) appendage and a base with three pores (circled) that lead to internal epithelium-lined crypts. The right half is a frontal section through the organ showing these epithelium-lined crypts (cr) containing the bacteria (stippled) within crypt diverticula and the relation of the crypts and ducts to the pores (circled), which open onto the surface. This section also shows the sinuses (arrows) within the appendages. The sinus spaces are continuous with the circulatory system and separate from the bacteria-containing crypts. ci, cilia; d, ducts; is, ink sac. Scale bar, 150 μm.

[1]Pacific Biomedical Research Center, Kewalo Marine Laboratory, University of Hawaii, 41 Ahui Street, Honolulu, HI 96813, USA. [2]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA. [3]Department of Microbiology, University of Iowa, Iowa City, IA 52242, USA. [4]Department of Microbiology, University of Georgia, Athens, GA 30602, USA. [5]Department of Medical Microbiology and Immunology, University of Wisconsin, 1300 University Avenue, Madison, WI 53706, USA.

*To whom correspondence should be addressed. E-mail: mjmcfallngai@wisc.edu

trigger inflammatory cytokine release, nitric oxide production, and organ injury in a rat model of bacterial sepsis (*12*).

Because mixed fragments of PGN with LPS could not consistently induce regression to the extent elicited by the intact symbiont (Fig. 2I), we reasoned that *V. fischeri* might have an active mechanism to release specific PGN fragments to the host, as opposed to the passive release due to cell lysis. Specific release of peptidoglycan by growing cells has only been observed in cultures of *Bordetella pertussis* (*13*) and *Neisseria gonorrhoeae* (*14*), both of which release large amounts of peptidoglycan monomers. The best studied of these fractions is tracheal cytotoxin (TCT; *N*-acetylglucosaminyl-1,6-anhydro-*N*-acetylmuramylalanyl-γ-glutamyldiaminopimelylalanine) (Fig. 3A), a disaccharide-tetrapeptide monomer of PGN, which causes the epithelial cytopathology of pertussis (*15*) and gonococcal infections (*16*). Fractionation of culture supernatants of *V. fischeri* by reversed-phase HPLC revealed a peak with an elution time identical to that of TCT. When purified, this fraction contained three amino acids: alanine, glutamic acid, and diaminopimelic acid,

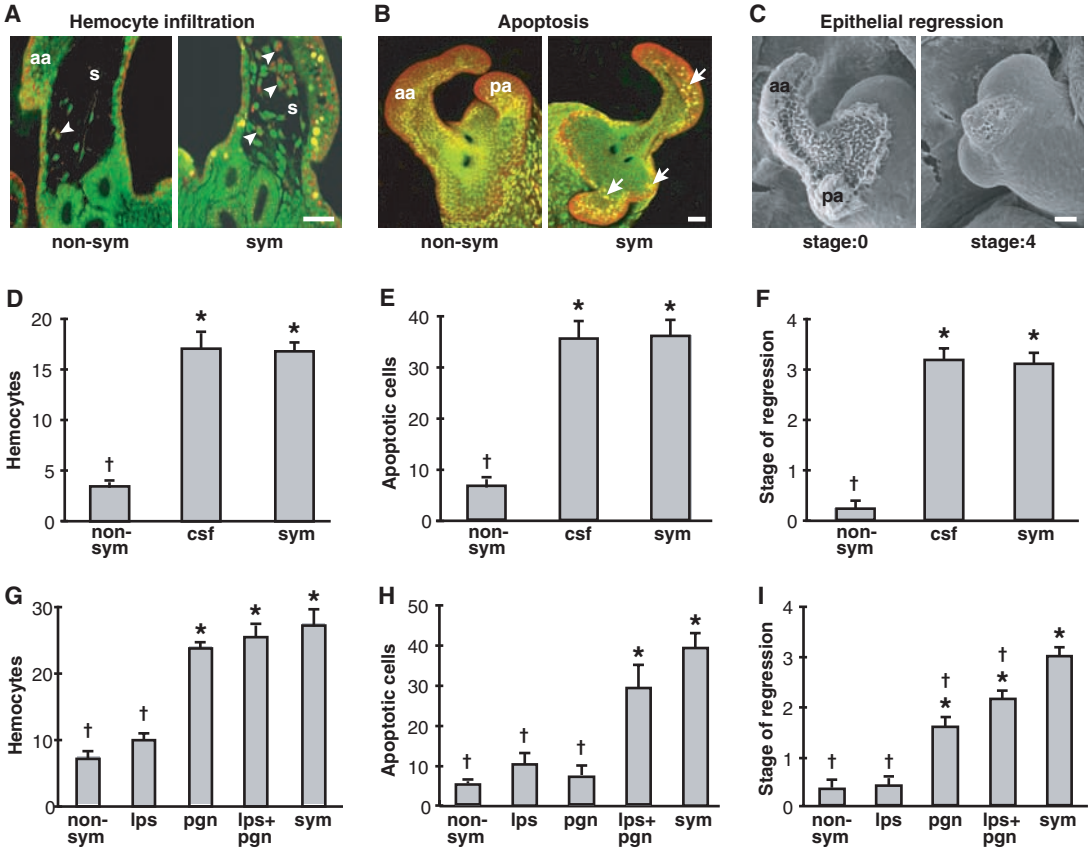


Fig. 2. Bacterial components induce light organ morphogenesis. (**A** and **B**) Confocal micrographs of nonsymbiotic [non-sym, i.e., uninfected (*25*)] and symbiotic (sym) organ epithelial fields stained with acridine orange (green) and Lysotracker (red). (A) Hemocytes (arrowheads) within the appendage sinuses (s). (B) Apoptotic cells, yellow foci (arrows). (**C**) SEMs of epithelial fields before (stage 0) and after (stage 4) regression. Scale bar, 50 μm. (**D** to **F**) Effects of *V. fischeri* cell surface fractions (csf) on hemocytes infiltration (D), induction of apoptosis (E), and epithelial regression (F). Animals were exposed to surface fractions at a protein concentration of 100 μg/ml of seawater. (**G** to **I**) Effects of *V. fischeri* surface components, LPS (10 μg/ml) and PGN (50 μg/ml), on hemocyte infiltration (G), induction of apoptosis (H), and epithelial regression (I). Data are means ± SEM for one of three replicate experiments (n = 8 to 12 per treatment). (*) indicates significant (P < 0.001) difference compared with non-sym. (†) indicates significant (P < 0.001) difference compared with sym.

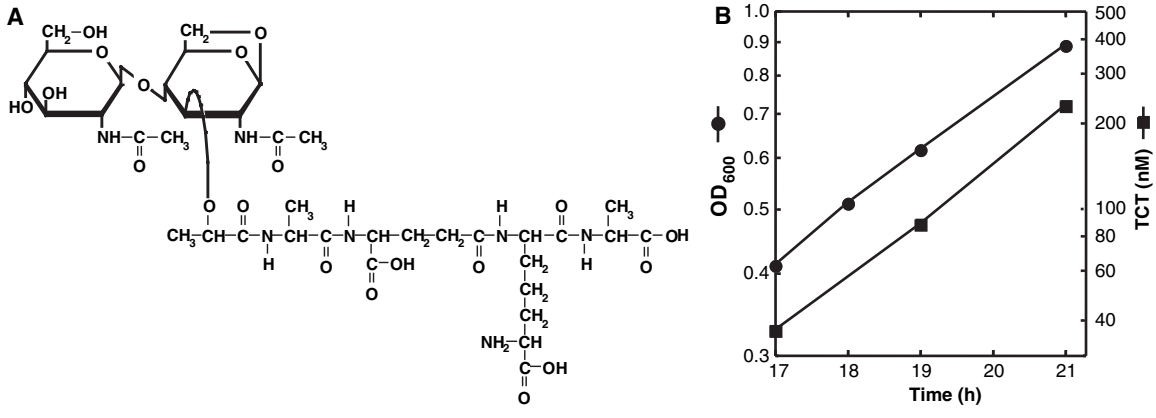


Fig. 3. TCT, a monomer of PGN, is released by growing cells of *V. fischeri*. (**A**) Structure of TCT. *V. fischeri* strain ES114 was cultured at room temperature in defined medium HM (*26*) supplemented with 2% glucose. A released fragment of *V. fischeri* PGN was determined to be identical to TCT by subjecting culture supernatants to solid phase extraction and two reverse-phase HPLC steps, as described for *B. pertussis* (*13*). A peak with an elution time corresponding to *B. pertussis* TCT was collected and further characterized by amino acid analysis and by matrix-assisted laser desorption mass spectrometry (*25*). (**B**) Production of TCT during a 4 hour period of log-phase growth. For sensitive quantification of TCT production, broth cultures were inoculated at OD_{600} = 0.04, after which log-phase *V. fischeri* culture supernatants were collected, subjected to solid phase extraction (*13*), and derivatized with phenylisothiocyanate. The resulting phenylthiocarbamyl (PTC) derivatives were separated by reversed-phase high-performance liquid chromatography using a C_8 column and detected at 254 nm. The amount of *V. fischeri* PTC-TCT in each sample was determined by comparing the peak area and elution time with an identically processed TCT standard. Results are representative of two experiments.

in molar proportions of 2:1:1. Mass spectrometry revealed a single species with a mass of 921 daltons. These data correspond precisely to TCT (*17*). Unlike most Gram-negative bacteria, *V. fischeri*, like *B. pertussis* (*18*), actively released appreciable amounts of TCT during log-phase growth (Fig. 3B).

TCT alone triggered hemocyte infiltration and regression of the epithelial fields at levels similar to those induced by intact *V. fischeri* (Fig. 4, A and C). TCT induced epithelial regression 4 days after exposures as brief as 14 hours (*9*), a time course similar to the intact symbiosis (*19*). The concentration of TCT required to trigger a detectable morphogenic response was as little as 10 nM, and the response was saturated at concentrations as low as 1 μM (0.9 μg/ml) (*9*). Comparatively, PGN, a multimer of TCT subunits, was generally not active below 50 μg/ml, and constant exposure was necessary to trigger epithelial regression by 4 days. To determine how specific the morphogenic responses were to TCT, we also tested muramyl dipeptide and glucosylmuramyl dipeptide, which are two smaller components of PGN known to signal host cell responses in various mammalian model systems (*20*, *21*). No morphogenic activity was detected in squid light organs when exposed to these components, either alone or in combination with LPS (*9*).

Under the conditions of the assay, TCT also triggered levels of apoptosis in the epithelial fields similar to those observed in the intact symbiosis (Fig. 4B), whereas LPS-free PGN did not induce apoptosis (Fig. 2H). Thus, the background level of LPS in the natural seawater, which typically ranges from 0.01 to 1.5 ng/ml (*22*), was sufficient to act together with the more potent TCT to induce the apoptotic effect (Fig. 4B). However, as cells of *V. fischeri* naturally colonize the crypts in high numbers and subsequently trigger morphogenesis from within these spaces (*19*), it is likely to be symbiont-specific LPS which works in concert with TCT to induce morphogenesis.

These results show that growing cells of *V. fischeri* release TCT, which acts as a potent morphogen to induce normal light organ morphogenesis in the squid host. Similar morphogenic effects of TCT have been reported for both *B. pertussis* and *N. gonorrhoeae*, which induce the loss of ciliated cells from mammalian respiratory and fallopian tube epithelia, respectively (*15*, *16*). In addition, TCT works in synergy with LPS to stimulate the production of inflammatory cytokines, nitric oxide, and the inhibition of DNA synthesis in hamster tracheal epithelium (*23*). In these studies, TCT-induced epithelial morphogenesis resulted from a direct interaction with the target epithelium. Note that, in the squid light organ, morphogenesis is triggered from within the crypt spaces, several cell layers away from the target epithelium (*19*). Thus, it will be interesting to identify and localize the TCT receptor(s) within the light organ and to decipher the host pathways that mediate these remote events.

These findings, as well as recent studies of mechanisms of tolerance to gut microbiota (*24*), demonstrate that "PAMPs" may be too narrow an acronym for eukaryotic-prokaryotic signaling associated with microbial molecules, such as LPS and PGN. The data suggest that a more general term, such as microbe associated molecular patterns (MAMPs), would be more appropriate to describe factors conserved and essential to the biology of microbes, which mediate recognition and response during host-microbe interactions.

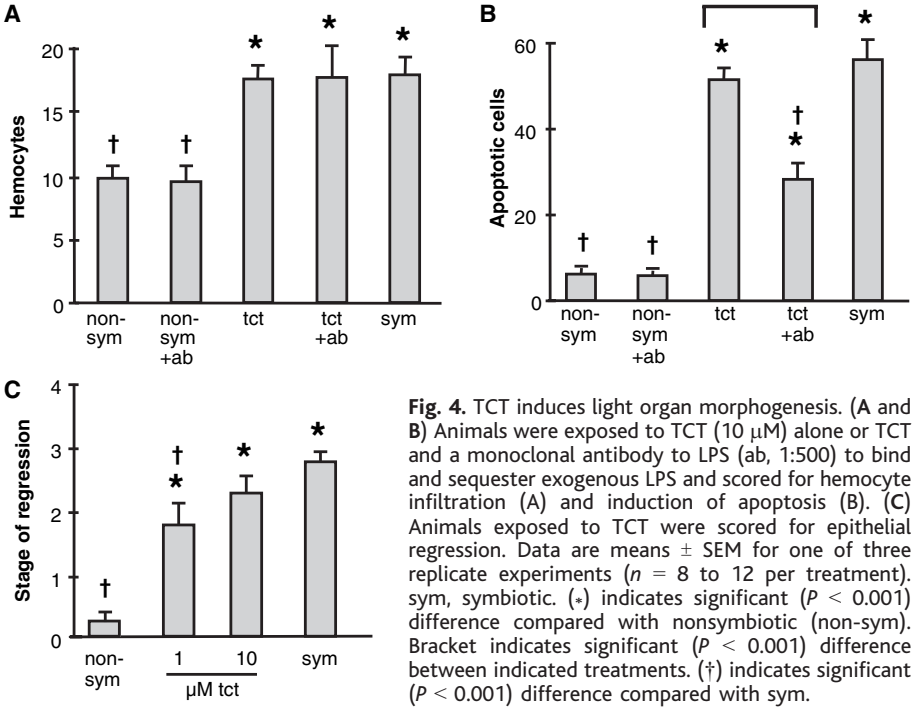


Fig. 4. TCT induces light organ morphogenesis. (**A** and **B**) Animals were exposed to TCT (10 μM) alone or TCT and a monoclonal antibody to LPS (ab, 1:500) to bind and sequester exogenous LPS and scored for hemocyte infiltration (A) and induction of apoptosis (B). (**C**) Animals exposed to TCT were scored for epithelial regression. Data are means ± SEM for one of three replicate experiments (n = 8 to 12 per treatment). sym, symbiotic. (*) indicates significant ($P < 0.001$) difference compared with nonsymbiotic (non-sym). Bracket indicates significant ($P < 0.001$) difference between indicated treatments. (†) indicates significant ($P < 0.001$) difference compared with sym.

References and Notes

1. R. Medzhitov, C. A. Janeway Jr., *Science* **296**, 298 (2002).
2. L. V. Hooper, *Trends Microbiol.* **12**, 129 (2004).
3. J. Xu, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10452 (2003).
4. M. J. McFall-Ngai, *Dev. Biol.* **242**, 1 (2002).
5. S. V. Nyholm, M. J. McFall-Ngai, *Nature Rev. Microbiol.*, **2**, 632 (2004).
6. S. V. Nyholm, E. V. Stabb, E. G. Ruby, M. J. McFall-Ngai, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10231 (2000).
7. S. V. Nyholm, B. Deplancke, H. R. Gaskins, M. A. Apicella, M. J. McFall-Ngai, *Appl. Environ. Microbiol.* **68**, 5113 (2002).
8. J. S. Foster, M. A. Apicella, M. J. McFall-Ngai, *Dev. Biol.* **226**, 242 (2000).
9. M. J. McFall-Ngai *et al.*, unpublished observations.
10. M. J. Pabst, S. Beranova-Giorgianni, J. M. Krueger, *Neuroimmunomodulation* **6**, 261 (1999).
11. Z. Q. Liu, G. M. Deng, S. Foster, A. Tarkowski, *Arthritis Res.* **3**, 375 (2001).
12. G. M. Wray, S. J. Foster, C. J. Hinds, C. Thiemermann, *Shock* **15**, 135 (2001).
13. B. T. Cookson, H. L. Cho, L. A. Herwaldt, W. E. Goldman, *Infect. Immun.* **57**, 2223 (1989).
14. R. S. Rosenthal, *Infect. Immun.* **24**, 869 (1979).
15. W. E. Goldman, D. G. Klapper, J. B. Baseman, *Infect. Immun.* **36**, 782 (1982).
16. M. A. Melly, Z. A. McGee, R. S. Rosenthal, *J. Infect. Dis.* **149**, 378 (1984).
17. B. T. Cookson, A. N. Tyler, W. E. Goldman, *Biochemistry* **28**, 1744 (1989).
18. R. S. Rosenthal, W. Nogami, B. T. Cookson, W. E. Goldman, W. J. Folkening, *Infect. Immun.* **55**, 2117 (1987).
19. J. A. Doino, M. J. McFall-Ngai, *Biol. Bull.* **189**, 347 (1995).
20. N. Inohara *et al.*, *J. Biol. Chem.* **278**, 5509 (2003).
21. S. Traub *et al.*, *J. Biol. Chem.* **279**, 8694 (2004).
22. D. M. Karl, F. C. Dobbs, in *Molecular Approaches to the Study of the Ocean*, K. E. Cooksey, Ed. (Chapman and Hall, London, 1998), pp. 29–89.
23. T. A. Flak, L. N. Heiss, J. T. Engle, W. E. Goldman, *Infect. Immun.* **68**, 1235 (2000).
24. S. Rakoff-Nahoum, J. Paglino, F. Eslami-Varzaneh, S. Edberg, R. Medzhitov, *Cell* **118**, 229 (2004).
25. Materials and methods are available on *Science* Online.
26. E. G. Ruby, K. H. Nealson, *Appl. Environ. Microbiol.* **34**, 164 (1977).

27. We are grateful to C. Chun, W. Crookes, M. Goodson, J. Graber, D. Millikan, E. Ruby, and A. Schaefer for critical reading of this manuscript; M. McMahon for technical assistance; and C. Yap for Fig. 1 illustrations. We also thank M. Hadfield for the provision of *P. luteoviolacea*. Supported by an NIH grant to M.M.-N., M.A.A., and E.V.S. (grant no. R01-AI50661), an NSF grant to M.M.-N. (grant no. IBN0211673), an NIH grant to E. G. Ruby and M.M.-N. (grant no. NCRR12294), a W. M. Keck Foundation grant to M.M.-N. and M.A.A., and a Canadian NSERC scholarship to T.A.K.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1186/DC1
Materials and Methods
References

1 July 2004; accepted 21 September 2004

Reversible Immunocontraception in Male Monkeys Immunized with Eppin

M. G. O'Rand,[1,2]* E. E. Widgren,[1,2] P. Sivashanmugam,[1,2]† R. T. Richardson,[1,2] S. H. Hall,[1,3] F. S. French,[1,3] C. A. VandeVoort,[4] S. G. Ramachandra,[5] V. Ramesh,[5] A. Jagannadha Rao[5]

Various forms of birth control have been developed for women; however, there are currently few options for men. The development of male contraceptives that are effective, safe, and reversible is desired for family planning throughout the world. We now report contraception of male nonhuman primates (*Macaca radiata*) immunized with Eppin, a testis/epididymis-specific protein. Seven out of nine males (78%) developed high titers to Eppin, and all of these high-titer monkeys were infertile. Five out of seven (71%) high–anti-Eppin titer males recovered fertility when immunization was stopped. This study demonstrates that effective and reversible male immunocontraception is an attainable goal. This method of immunocontraception may be extended to humans.

Although several different choices and approaches are available for contraception in women, the choices for men are currently limited to condoms and vasectomy (*1*, *2*). Male hormonal contraceptives (*3*, *4*) developed over the past several years have now advanced to clinical trials, and the outcome of these studies may determine whether the suppression of sperm production through androgen regulation can become a realistic product. Immunocontraception, an alternative nonhormonal method, has been studied for many years (*1*, *5*), with the major emphasis on immunization of females to prevent pregnancy (*6*) or fertilization (*7*). In the present study, we report the successful contraception of male nonhuman primates (*M. radiata*) immunized with Eppin, a testis/epididymis-specific protein (*8*, *9*). This represents a non–hormonally disruptive male immunocontraceptive for primates.

Before using the monkey as a model for the test of a male immunocontraceptive, we determined the presence of both Eppin and immunoglobulin (IgG) in the male reproductive tract, the immunogenicity of Eppin, and the effects of immunization on sperm motility at the University of California, Davis (UC-Davis). Results from these studies are shown in figs. S1 to S3 and the supporting online material text (*10*). The studies on normal *Macaca* monkeys at UC-Davis allowed us to determine that IgG is present in the normal epididymal tract. Moreover, immunization of two monkeys at UC-Davis indicated that the immune response to recombinant Eppin could influence the titer of antibodies to Eppin in their semen. These results are similar to a previous study in female macaques demonstrating that different serum titers resulted in correspondingly different antibody titers in oviductal fluid (*11*). Consequently, in the fertility study described below, two male monkeys in the initial immune group were dropped because they could not sustain a high serum titer and were unlikely to have a high semen titer. The lack of a strong immune response to an immunogen in a particular individual animal is a reflection of the major histocompatibility complex and T cell response (*12*) as well as the antigen's availability to regulate the immune response (*13*). Such responses are found in heterozygous populations and would need further study before proceeding with additional Eppin fertility trials, which might include a linear B cell epitope (fig. S7).

We tested the effect of Eppin immunization on male fertility at the Indian Institute of Science. Six adult male monkeys (*M. radiata*) were immunized with human recombinant Eppin and six controls received adjuvant only. High–anti-Eppin titers were detected in four of the six monkeys immunized with Eppin in squalene; two monkeys were low responders with titers <1:400 on postimmunization day

Table 1. Fertility test of Eppin-immunized male monkeys (*M. radiata*). The seven males were 7 to 12 years old and had each sired 1 to 3 offspring in the previous 1 to 2 years. No pregnancies resulted from immunized males. Each exposed female had an estradiol 17β surge, indicating that an ovulation probably occurred in that cycle. ELISA O.D., enzyme-linked immunosorbent assay O.D. at 1:1000 on nearest day of cohabitation.

Male monkey no.	Cycle no.	Female monkey no.	ELISA O.D.	Days after first immunization
		Group 1 Eppin (squalene)		
602				
	I	55	0.39	476
	II	88	1.19	522
	III	107	1.19	528
619				
	I	84	0.67	475
	II	64	0.56	523
	III	23	0.56	530
625				
	I	60	1.48	397
	II	97	1.00	525
	III	89	1.15	568
679				
	I	19	0.70	475
	II	47	0.62	534
	III	82	0.62	542
		Group 2 Eppin (CFA)		
610				
	I	29	0.91	147
	II	73	0.59	225
	III	97	1.19	250
656				
	I	107	1.86	147
	II	47	1.93	225
	III	55	2.14	250
657				
	I	79	1.27	147
	II	6	1.53	276
	III	82	1.00	323

[1]Laboratories for Reproductive Biology, [2]Department of Cell and Developmental Biology, and [3]Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [4]California National Primate Research Center, University of California, Davis, CA 95616, USA. [5]Department of Biochemistry, Primate Research Laboratory, Indian Institute of Science, Bangalore 560012, India.

*To whom correspondence should be addressed. E-mail: morand@unc.edu

†Present address: Department of Urology, Duke University, Durham, NC 27710, USA.

Table 2. Fertility test of male monkeys (*M. radiata*) in adjuvant control group. The six males were 10 to 11 years old and had each sired 1 to 3 offspring in the previous 1 to 3 years. Male 609 impregnated two different females. Four out of six males (67%) impregnated females.

Male monkey no.	Female monkey no.	Days after first immunization when conceived
569	37, 107, 47	No pregnancy
604	35	650
607	87	650
609	86; 6	502; 594
688	96	502
690	6, 84, 11	No pregnancy

126. Because the purpose of this study was to test the efficacy of antibodies to Eppin on fertility, the two low-titer monkeys were dropped from the study group and their fertility was not tested. Three additional males were added, which were immunized (primary immunization) with recombinant human Eppin in complete Freund's adjuvant (CFA) to boost immunogenicity. The original four monkeys immunized with Eppin in squalene were designated group 1, and the three additional males immunized with Eppin in CFA were designated group 2.

Antibody titers in all of the monkeys were >1:10,000 at the time the matings started and remained elevated throughout the mating period. Figure S4, A (group 1) and B (group 2), shows the mean optical density (O.D.) value at 450 nm for the monkeys at a 1:1000 dilution of serum. A titer of >1:1000 was sustained for 775 days in group 1 (fig. S4A) and for 481 days in group 2 (fig. S4B). There was no effect on serum testosterone levels in the immunized males in either group 1 or group 2 compared with control values (fig. S5) and no effect on sperm counts in either group (fig. S6).

Each male monkey in the immune and control groups was subjected to fertility testing by cohabiting with a proven fertile female between days 9 and 14 of her menstrual cycle. Each male was exposed to three ovulatory cycles of three different females to test their fertility. Immune and control groups began fertility testing on days 390 to 397 (Table 1 and table S2; for group 2, day 390 is 147 days after their first day of immunization). Group 1 completed testing on day 568, group 2 completed testing on day 566 (323 days after their first day of immunization), and the control group completed testing on day 691 (table S2). None of the immunized monkeys was able to impregnate females, indicating that males with sustained high–anti-Eppin titers were infertile (Table 1). Four monkeys in the adjuvant control group impregnated 5 females (4 out of 6, 67%, Table 2). All the monkeys used for breeding exhibited ovulatory cycles (table S1).

After the completion of fertility testing, immunizations of monkeys stopped on day 691 (day 448 of immunization for group 2). Groups 1 and 2 were maintained without further immunizations for 450 days (group 1; Eppin/squalene) and 451 days (group 2; Eppin/CFA), respectively, to test their ability to recover fertility after immunization. During this recovery time period, three of four monkeys in the Eppin/squalene group and two of three monkeys in the Eppin/CFA group recovered their fertility for a total recovery of 71% (5 out of 7; table S3). The males exhibited no symptoms of autoimmune disease and had no detectable serum titer of antibody to Eppin at 1:1000 dilutions.

This study demonstrates that effective and reversible male immunocontraception in primates is an attainable goal. We found that a high serum titer (>1:1000), sustained over several months, achieves an effective level of contraception. Seven out of nine males (78%) developed high titers to Eppin, and all these high-titer monkeys were infertile. Five out of seven (71%) high–anti-Eppin titer males recovered fertility when immunization was stopped.

Eppin on the surface of spermatozoa and in semen is bound to semenogelin (*14*), which is involved in coagulum formation in the ejaculate. We can speculate that one mechanism to explain the infertility is that antibodies to Eppin interfere with normal Eppin interaction with the sperm surface and with semenogelin.

References and Notes

1. S. J. Nass, J. F. Strauss, Eds., *New Frontiers in Contraceptive Research* (National Academies Press, Washington, DC, 2004).
2. C. Holden, *Science* **296**, 2172 (2002).
3. A. Kamischke, E. Nieschlag, *Trends Pharmacol. Sci.* **25**, 49 (2004).
4. J. K. Amory, W. J. Bremner, *Trends Endocrinol. Metab.* **11**, 61 (2000).
5. M. G. O'Rand, I. A. Lea, *J. Reprod. Immunol.* **36**, 51 (1997).
6. G. P. Talwar *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8532 (1994).
7. P. Primakoff, W. Lathrop, L. Woolman, A. Cowan, D. Myles, *Nature* **335**, 543 (1988).
8. R. T. Richardson *et al.*, *Gene* **270**, 93 (2001).
9. P. Sivashanmugam *et al.*, *Gene* **312**, 125 (2003).
10. Materials and methods are available as supporting material on *Science* Online.
11. I. A. Lea, B. Kurth, M. G. O'Rand, *Biol. Reprod.* **58**, 794 (1998).
12. I. A. Lea *et al.*, *Biol. Reprod.* **59**, 527 (1998).
13. R. M. Zinkernagel, H. Hengartner, *Science* **293**, 251 (2001).
14. R. T. Richardson, E. Widgren, Z. Wang, P. Sivashanmugam, M. G. O'Rand, *Biol. Reprod. Suppl.* **70**, 98 (abstr.) (2004).

15. Supported by grant CIG-96-06 from the Consortium for Industrial Collaboration in Contraceptive Research Program of Contraception Research and Development (CONRAD).

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1189/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 to S3
References

29 April 2004; accepted 17 September 2004

A Cluster of Metabolic Defects Caused by Mutation in a Mitochondrial tRNA

Frederick H. Wilson,[1,2,3*] Ali Hariri,[1,4*] Anita Farhi,[1,2] Hongyu Zhao,[2,5] Kitt Falk Petersen,[4] Hakan R. Toka,[1,2] Carol Nelson-Williams,[1,2] Khalid M. Raja,[8] Michael Kashgarian,[6] Gerald I. Shulman,[1,4,7] Steven J. Scheinman,[8] Richard P. Lifton[1,2,3,4†]

Hypertension and dyslipidemia are risk factors for atherosclerosis and occur together more often than expected by chance. Although this clustering suggests shared causation, unifying factors remain unknown. We describe a large kindred with a syndrome including hypertension, hypercholesterolemia, and hypomagnesemia. Each phenotype is transmitted on the maternal lineage with a pattern indicating mitochondrial inheritance. Analysis of the mitochondrial genome of the maternal lineage identified a homoplasmic mutation substituting cytidine for uridine immediately 5′ to the mitochondrial transfer RNA^{Ile} anticodon. Uridine at this position is nearly invariate among transfer RNAs because of its role in stabilizing the anticodon loop. Given the known loss of mitochondrial function with aging, these findings may have implications for the common clustering of these metabolic disorders.

Hypertension and dyslipidemia are important risk factors for many common cardiovascular diseases, including myocardial infarction, stroke, and congestive heart failure (*1*, *2*). These traits are concordant in individual patients more often than expected by chance (*3*, *4*). Large epidemiologic studies have demonstrated that subjects with hypertension

have a marked increase in the prevalence of hypercholesterolemia, hypertriglyceridemia, hypomagnesemia, diabetes, insulin resistance, and obesity (*5*–*9*). Various combinations of these abnormalities affect up to a quarter of the U.S. adult population and are referred to as the metabolic syndrome, syndrome X, or dyslipidemic hypertension. The factors accounting for this phenotypic clustering are unknown, although obesity, insulin resistance, and increased local glucocorticoid tone have been suggested to play a role (*4*, *10*). Although rare mutations with large effects on blood pressure (*11*), lipids (*12*), insulin resistance (*13*, *14*), obesity (*15*), and magnesium (*16*) have established critical pathways for homeostasis of each of these traits, they have typically affected only one of these phenotypes and therefore have not provided an explanation for their clustering (*17*).

A Caucasian kindred (K129) was ascertained through a proband with hypomagnesemia. Evaluation of her extended kindred revealed a high prevalence of hypomagnesemia, hypertension, and hypercholesterolemia. We ultimately performed a detailed clinical evaluation of 142 blood relatives in the kindred (Fig. 1). Including the index case, 38 members had hypertension (with blood pressure > 140/90 mm Hg or on treatment for hypertension), 33 had hypercholesterolemia (with total cholesterol > 200 mg/dl or on treatment for hypercholesterolemia), and 32 had clinically significant hypomagnesemia (range 0.8 to 1.7 mg/dl, normal 1.8 to 2.5 mg/dl).

Because it is the least common of these traits in the general population, we initially focused on the distribution of hypomagnesemia in the kindred. Hypomagnesemic individuals are distributed through four generations and 16 sibships, and both genders are affected (Fig. 1); there was no significant effect of age on Mg^{2+} levels and no hypomagnesemic subjects were taking Mg^{2+}-altering medications. All 32 members with hypomagnesemia are on the same maternal lineage (Figs. 1 and 2A). Affected fathers never transmitted the trait to their offspring (0 of 17 offspring), whereas affected mothers transmitted the trait to a high fraction of their offspring (16 of 21). These features are hallmarks of inheritance via the mitochondrial genome. The probability of all 32 hypomagnesemic subjects being on the maternal lineage by chance is extremely small (χ^2 = 49, $P < 10^{-11}$), strongly supporting mitochondrial transmission. Autosomal dominant transmission with imprinting was much less likely from the observed distribution (the odds favoring mitochondrial transmission were $>10^6$:1) (*18*). A genome-wide analysis of linkage was performed and found no evidence for a shared segment of the nuclear genome among hypomagnesemic subjects (*18*).

Quantitative serum Mg^{2+} levels were lower in individuals from the maternal lineage compared with relatives in the nonmaternal lineage (Fig. 2A) ($P = 2 \times 10^{-9}$). Members of the maternal lineage had a marked increase in the urinary fractional excretion of Mg^{2+} (Fig. 2B) ($P = 0.0001$); this effect was most pronounced among subjects with hypomagnesemia (Fig. 2B) ($P = 5 \times 10^{-6}$ comparing hypomagnesemic subjects versus subjects in the nonmaternal lineage), establishing impaired renal Mg^{2+} reabsorption as the cause of hypomagnesemia in K129. Evaluation of other urinary electrolytes was notable for reduced urinary calcium on the maternal lineage (Fig. 2C) ($P = 0.0005$), again predominantly among hypomagnesemic subjects ($P = 4 \times 10^{-6}$) despite normal serum calcium levels. Hypomagnesemia with reduced urinary calcium is characteristic of a primary defect in the renal distal convoluted tubule (DCT) (*16*). In addition, hypokalemia due to inappropriate renal loss was seen more frequently on the maternal lineage (Fig. 2D) ($\chi^2 = 11.6$, $P = 0.0007$), predominantly among hypomagnesemic subjects. There was no difference in 24-hour urinary sodium excretion between maternal and nonmaternal lineages. Electrolyte values are summarized in table S1.

Hypertension also segregated with the maternal lineage. Thirty of 53 adults on the maternal lineage had blood pressure greater than 140/90 mm Hg or were being treated with antihypertensive medication versus 8 of 53 on the nonmaternal lineages ($\chi^2 = 19.9$, $P < 0.00001$). The prevalence of hypertension on the maternal lineage showed a marked age dependence, increasing from 5% in subjects under age 30 (1 of 20 subjects), to 44% in those from age 30 to 50 (10 of 23 subjects), and to 95% in those over age 50 (19 of 20 subjects). Because the oldest generations of K129 are enriched for the ma-

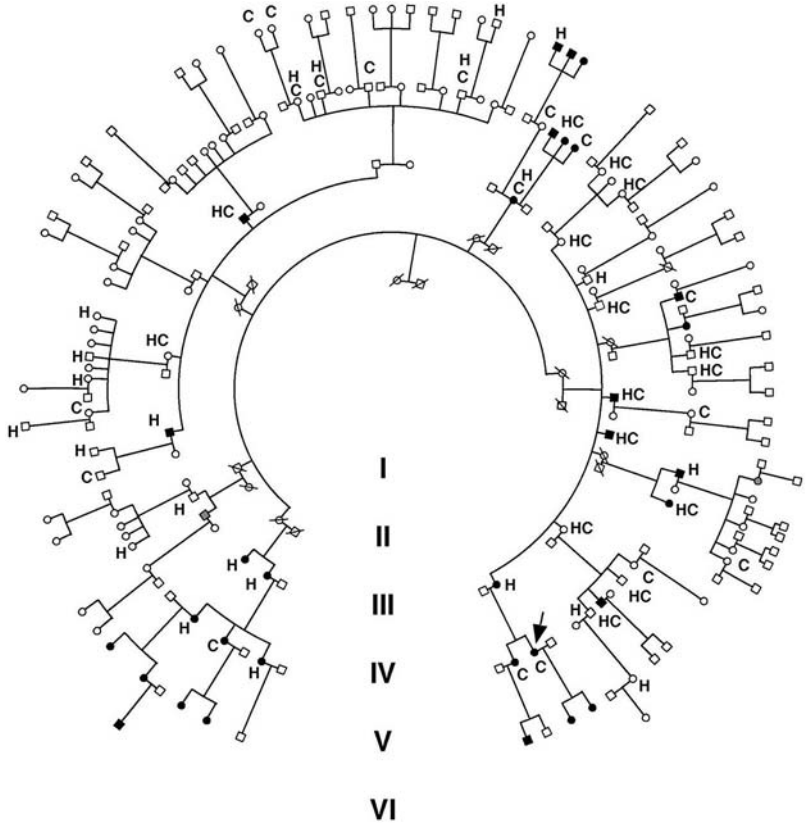


Fig. 1. The structure of Kindred 129. Individuals with serum $Mg^{2+} < 1.8$ mg/dl are indicated by black symbols. Family members taking antihypertensive medications or having blood pressures over 140/90 mm Hg are indicated by an H. Members with hypercholesterolemia (serum cholesterol > 200 mg/dl or taking lipid-lowering agents) are denoted by C. Blood relatives who did not have electrolyte values measured are indicated by gray symbols. The index case is indicated by an arrow.

[1]Howard Hughes Medical Institute, [2]Department of Genetics, [3]Department of Molecular Biophysics and Biochemistry, [4]Department of Internal Medicine, [5]Department of Biostatistics, [6]Department of Pathology, [7]Department of Cell and Molecular Physiology, Yale University School of Medicine, New Haven, CT 06510, USA. [8]Department of Medicine, State University of New York Upstate Medical University, Syracuse, NY 13210, USA.

*These authors contributed equally to this manuscript. †To whom correspondence should be addressed. E-mail: richard.lifton@yale.edu

ternal lineage (Fig. 1), we reanalyzed the data, excluding subjects over age 60; the results remain highly significant ($P = 0.0005$) (supporting online text). The prevalence of hypertension on the maternal lineage is also high compared with the general population (supporting online text).

Quantitative assessment confirmed the effect of maternal lineage on blood pressure (Fig. 3, A and B, and Table 1). After adjustment of blood pressure for the major covariates age, sex, and body mass index (BMI), adults on the maternal lineage had highly significant increases in systolic and diastolic blood pressures compared with their nonmaternal relatives. Among adults age 18 to 60, maternal lineage increased systolic blood pressure by an average of 13 mm Hg ($P = 0.00007$) and diastolic blood pressure by 5 mm Hg ($P = 0.002$). Similar results are seen in analysis of all adults. Estimates of these quantitative effects are conservative because of the higher use of antihypertensive medication among members of the maternal lineage. Plasma renin and aldosterone levels were no different between members of the maternal and nonmaternal lineages (table S2).

Hypercholesterolemia also segregated with the maternal lineage. Twenty-four of 46 adults on the maternal lineage had fasting total cholesterol of >200 mg/dl or were being treated with cholesterol-lowering medication versus 9 of 49 on the nonmaternal lineage ($\chi^2 = 12.4$, $P = 0.0004$). The relationship remained highly significant ($P = 0.0008$) when the analysis was restricted to adults age 18 to 60. Similar results were obtained for elevated fasting low-density lipoprotein (LDL) cholesterol (LDL > 130 mg/dl; $\chi^2 = 11.6$, $P = 0.0007$). Quantitative analysis of total cholesterol among adults age 18 to 60 after adjustment for age, sex, and BMI revealed that maternal lineage increased total cholesterol by an average of 26 mg/dl (Fig. 3C and Table 1). This increase is attributable to elevations in LDL and very low-density lipoprotein (VLDL), with no effect on fasting high-density lipoprotein (HDL) or triglycerides (Fig. 3D, Table 1, and fig. S1). Similar results are seen among all adult subjects. The magnitude of these effects is likely an underestimate because of the increased use of cholesterol-lowering agents among maternal relatives.

In sum, of 45 adults on the maternal lineage who had all three traits measured, 38 had one or more of hypertension, hypercholesterolemia, or hypomagnesemia, 26 had two or more, and 7 had all three (fig. S2). The maternal lineage accounts for virtually all of the clustering of these traits in K129 (fig. S2).

Collectively, these data provide strong evidence for a mitochondrial mutation as the cause of the syndrome in K129. Because Southern blotting revealed no evidence of mitochondrial deletion, we performed direct sequencing and single-strand conformational polymorphism analysis of the entire mitochondrial genome to search for sequence variants. Fourteen variants were identified on the maternal lineage; 13 are previously identified polymorphisms of no known consequence (table S3). One variant, however, is a previously undescribed thymidine-to-cytidine transition at nucleotide 4291, which lies within the mitochondrial $tRNA^{Ile}$ gene (GenBank accession no. NC_001807) (Fig. 4, A and B). This mutation is found only on the maternal lineage in K129, does not appear among the thousands of mitochondrial genomes previously sequenced (*19*), and was absent

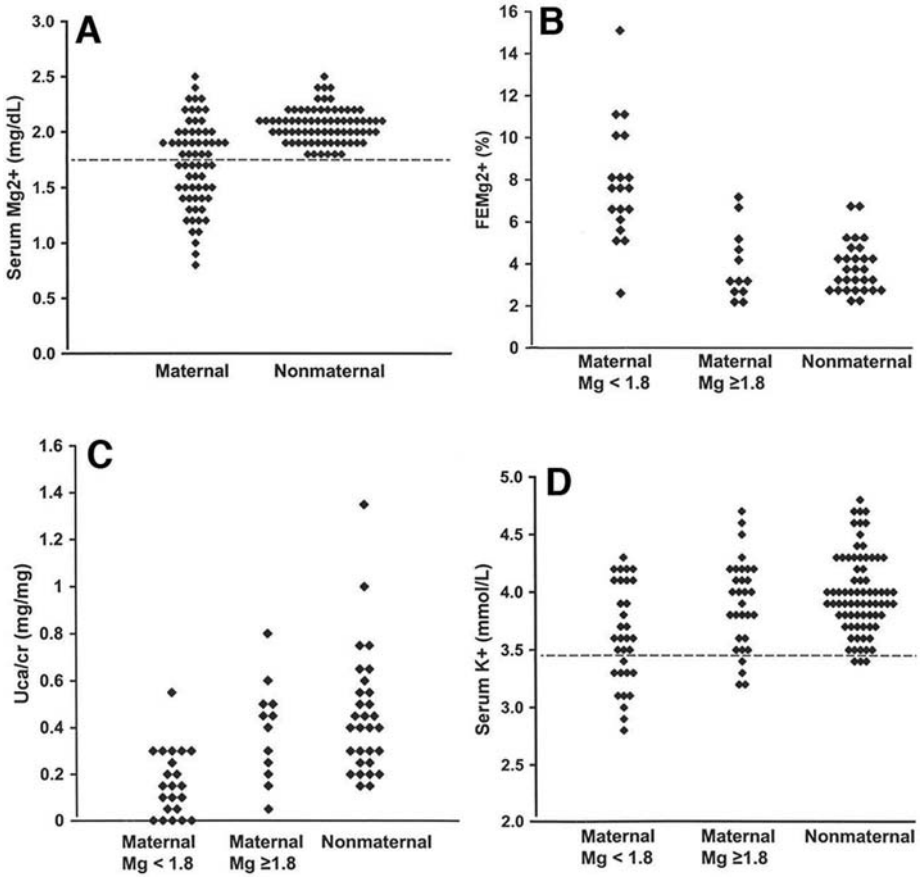


Fig. 2. Renal hypomagnesemia, hypocalciuria, and hypokalemia in the maternal lineage of K129. (**A**) Serum Mg^{2+} values for individuals in maternal and nonmaternal lineages of K129 are shown and are significantly different ($P = 2 \times 10^{-9}$). (**B**) Fractional renal Mg^{2+} excretion ($FEMg^{2+}$) on the maternal and nonmaternal lineages is shown; on the maternal lineage, individuals with normal and low Mg^{2+} levels are separated. Hypomagnesemic subjects have significantly elevated fractional excretion of Mg^{2+}, indicating a renal defect ($P = 0.0001$ comparing maternal to nonmaternal; $P = 5 \times 10^{-6}$ comparing hypomagnesemic subjects versus those not in the maternal lineage). (**C**) Urinary calcium to creatinine ratios (UCa/cr) are shown grouped as in (B); maternal subjects have significantly reduced urinary calcium levels ($P = 0.0005$). (**D**) Serum K^+ levels. Hypokalemia is seen predominantly on the maternal lineage among hypomagnesemic subjects.

Table 1. Age, sex, and BMI-adjusted traits in adults age 18 to 60 in maternal and nonmaternal lineages of K129. Values are mean ± SEM. Total cholesterol, LDL, VLDL, HDL, triglyceride, glucose, and insulin sensitivity (*18*) were measured after an overnight fast. SBP, systolic blood pressure; DBP, diastolic blood pressure; HOMA, homeostasis model assessment.

	Nonmaternal	Maternal	*P*
SBP (mm Hg)	122 ± 2	135 ± 3	0.00007
DBP (mm Hg)	77 ± 1	82 ± 1	0.002
Total cholesterol (mg/dl)	173 ± 4	199 ± 7	0.002
LDL + VLDL (mg/dl)	124 ± 5	150 ± 8	0.004
HDL (mg/dl)	50 ± 2	49 ± 3	0.78
Triglyceride (mg/dl)	129 ± 14	148 ± 22	0.46
Glucose (mg/dl)	84 ± 1	95 ± 11	0.28
HOMA	3.5 ± 0.3	4.0 ± 0.3	0.22

among 170 unrelated control individuals. Polymerase chain reaction–restriction fragment length polymorphism analysis revealed that this mutation is apparently homoplasmic in leukocytes of all members of the maternal lineage regardless of phenotype, with the assay sufficiently sensitive to detect 1% heteroplasmy (fig. S3) (*18*).

The thymidine-to-cytidine mutation in K129 occurs immediately 5′ to the $tRNA^{Ile}$ anticodon (Fig. 4C). Uridine at this position is one of the most extraordinarily conserved bases in the biological world. It is conserved in every sequenced isoleucine tRNA, including 242 different species of archaebacteria, eubacteria, unicellular and multicellular eukaryotes, animals, plants, chloroplasts, and mitochondria (*20*). Moreover, uridine is conserved at this position in virtually all sequenced tRNAs of all specificities (96% of 4300 tRNAs among all species); nearly all exceptions are eukaryotic initiator $tRNA^{Met}$ genes (*20*). The extreme conservation of uridine at this position is explained by the structure of tRNAs. The anticodon loop results from a sharp turn in the phosphodiester backbone, allowing presentation of the anticodon to its cognate mRNA codon in the ribosome (*21*, *22*). This turn is stabilized by a hydrogen bond between the amino group of the conserved uridine and the phosphate backbone of the third base of the anticodon (*22*, *23*). Cytidine lacks this amino group and cannot form this hydrogen bond. Biochemical studies with anticodon stem-loop analogs of tRNAs have been performed and indicate that substitution of cytidine for uridine at this position markedly impairs ribosome binding (*23*), providing evidence of the functional importance of this mutation.

Members of K129 were carefully evaluated for the presence of additional clinical phenotypes commonly associated with mitochondrial dysfunction. The prevalence of migraine headache, sensorineural hearing loss, and hypertrophic cardiomyopathy were increased on the maternal lineage (supporting online text). Measures of fasting HDL, triglycerides, insulin resistance, BMI, and diabetes mellitus were not significantly different between the two lineages (Table 1 and figs. S1 and S4).

Immunohistochemistry of a skeletal muscle biopsy from a member of the maternal lineage revealed an increase in ragged red fibers and subsarcolemmal succinate dehydrogenase staining, characteristic features of individuals carrying mitochondrial mutations (fig. S5, A and B) (*24*). Electron microscopy of the biopsy demonstrated cytoplasmic lipid accumulation, increased glycogen stores, and dysmorphic mitochondrial cristae, further signs of mitochondrial dysfunction (fig. S5C). Finally, in vivo nuclear magnetic resonance

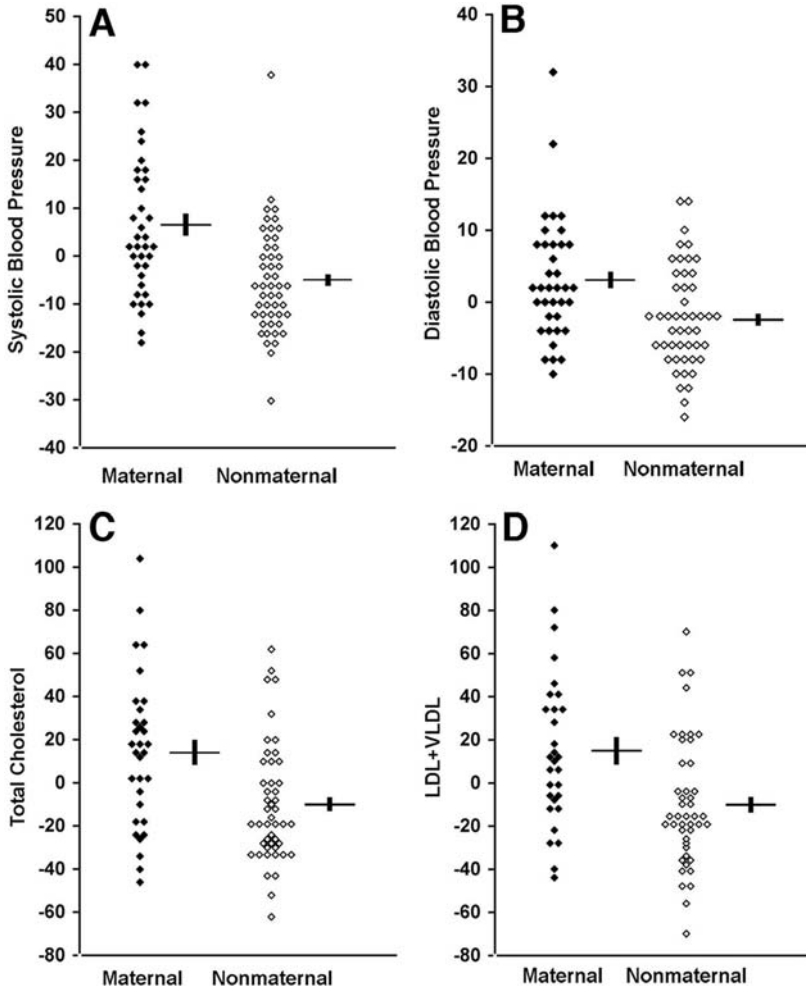


Fig. 3. Quantitative blood pressure and cholesterol values in K129. Values in maternal and nonmaternal K129 members between the ages of 18 and 60 are shown. All values represent difference from the mean value after adjustment for age, sex, and BMI. For blood pressure and lipids, units are mm Hg and mg/dl, respectively. Mean and SEM values are indicated for maternal and nonmaternal groups. Values are significantly elevated in members of the maternal lineage. (**A**) Systolic blood pressure ($P = 0.00007$). (**B**) Diastolic blood pressure ($P = 0.002$). (**C**) Fasting total cholesterol ($P = 0.002$). (**D**) Fasting LDL + VLDL cholesterol ($P = 0.004$).

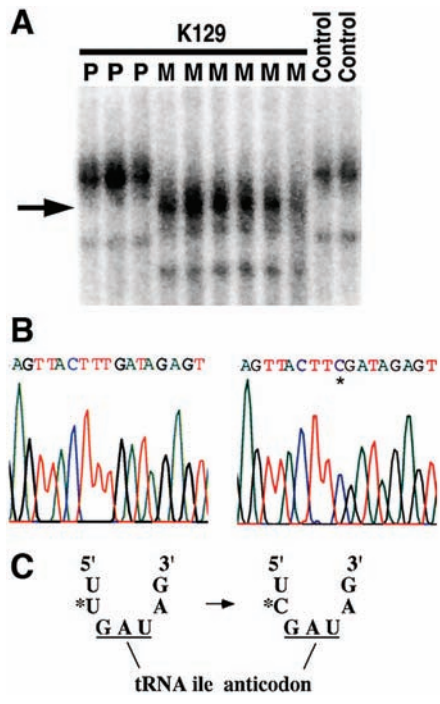


Fig. 4. Mitochondrial $tRNA^{Ile}$ mutation in K129. Mitochondrial DNA from both blood leukocytes and renal epithelial cells was analyzed and yielded identical results. (**A**) A fragment of mtDNA containing the $tRNA^{Ile}$ gene was amplified from members of K129 and normal controls and was fractionated by nondenaturing gel electrophoresis (*18*). A thymidine-to-cytidine variant (indicated by arrow) is present in individuals from the maternal lineage (M) but absent in offspring of affected males (paternal lineage, P) and unrelated controls. (**B**) The sequence of a portion of the mitochondrial $tRNA^{Ile}$ gene from the amplicon in (A) from a wild-type control (left) and a member of the maternal lineage of K129 (right). A single base substitution (asterisk) changes the wild-type thymidine to cytidine. (**C**) The T → C transition alters the nucleotide immediately 5′ to the $tRNA^{Ile}$ anticodon.

(NMR) spectroscopy of skeletal muscle in this patient demonstrated normal tricarboxylic acid cycle flux but reduced adenosine triphosphate (ATP) production, suggesting impaired coupling of these processes (fig. S5, D and E). Additional studies of other kindred members will be required to establish the frequency and severity of these manifestations.

These findings establish a causal relationship between a mitochondrial mutation and hypertension, hypercholesterolemia, and hypomagnesemia. The mitochondrial origin of this disorder is of particular interest given recent evidence implicating mitochondrial dysfunction in type 2 diabetes mellitus and insulin resistance, other components of the metabolic syndrome. Rare mitochondrial mutations cause diabetes with deafness (*25*). In vivo NMR of skeletal muscle has linked loss of mitochondrial function to insulin resistance (*26*). Finally, expression of genes involved in oxidative phosphorylation is reduced among patients with type 2 diabetes mellitus and insulin resistance (*27*). Thus, although insulin resistance, obesity, and hypertriglyceridemia are absent in K129, these traits have been previously linked to loss of mitochondrial function. These observations raise the possibility that all the features of the metabolic syndrome can result from pleiotropic effects of impaired mitochondrial function; we speculate that the loss of mitochondrial function with aging (*26*, *28*) might commonly contribute to all components of the metabolic syndrome.

The variation in the phenotypic consequences of this homoplasmic mitochondrial mutation is notable. Hypomagnesemia, hypertension, and hypercholesterolemia each show ~50% penetrance among adults on the maternal lineage. Incomplete penetrance arising from homoplasmic mutations is well described and has been attributed to nuclear genome and/or environmental modifiers (*29*). The nearly stochastic distributions of these traits on the maternal lineage (fig. S2) and the nonsignificant correlations among their quantitative values on the maternal lineage suggests that these are independent, pleiotropic effects of the mitochondrial mutation.

Prior studies suggest potential mechanisms linking each trait to impaired mitochondrial function. Cells of the DCT have the highest energy consumption of the nephron (*30*), and Mg^{2+} reabsorption in the DCT requires ATP-dependent Na^{+} reabsorption (*31*). Inhibitors of mitochondrial ATP production increase cholesterol biosynthesis while inhibiting clearance in vitro (*32*). Finally, reduced ATP production has been reported in animal models of hypertension (*33*). Further work will be required to elucidate the molecular mechanisms linking genotype and phenotype.

The results of this study suggest that the loss of mitochondrial function with age (*26*, *28*) could contribute to the characteristic age-related increase in blood pressure (*34*) and to its clustering with hypocholesterolemia in the general population. The mutation in K129 results in a complex pattern of phenotypic clustering that is reminiscent of the frequent but not obligatory clustering seen in the general population. This highlights the complexity that can arise from a single mutation because of the combined effects of reduced penetrance and pleiotropy and underscores the value of studying very large kindreds. The present findings motivate further investigation of a potential role for mitochondrial dysfunction in common forms of hypertension and hypercholesterolemia.

References and Notes

1. J. Stamler, D. Wentworth, J. D. Neaton, *JAMA* **256**, 2823 (1986).
2. A. Mosterd *et al.*, *N. Engl. J. Med.* **340**, 1221 (1999).
3. D. L. Wingard, E. Barrett-Connor, M. H. Criqui, L. Suarez, *Am. J. Epidemiol.* **117**, 19 (1983).
4. G. M. Reaven, *Diabetes* **37**, 1595 (1988).
5. M. H. Criqui *et al.*, *Circulation* **73**, I40 (1986).
6. R. R. Williams *et al.*, *JAMA* **259**, 3579 (1988).
7. S. Mizushima, F. P. Cappuccio, R. Nichols, P. Elliott, *J. Hum. Hypertens.* **12**, 447 (1998).
8. J. M. Peacock, A. R. Folsom, D. K. Arnett, J. H. Eckfeldt, M. Szklo, *Ann. Epidemiol.* **9**, 159 (1999).
9. F. Guerrero-Romero, M. Rodriguez-Moran, *Acta Diabetol.* **39**, 209 (2002).
10. H. Masuzaki *et al.*, *Science* **294**, 2166 (2001).
11. R. P. Lifton, A. G. Gharavi, D. S. Geller, *Cell* **104**, 545 (2001).
12. J. L. Goldstein, M. S. Brown, *Science* **292**, 1310 (2001).
13. G. I. Bell, K. S. Polonsky, *Nature* **414**, 788 (2001).
14. S. George *et al.*, *Science* **304**, 1325 (2004).
15. S. O'Rahilly, I. S. Farooqi, G. S. Yeo, B. G. Challis, *Endocrinology* **144**, 3757 (2003).
16. M. Konrad, K. P. Schlingmann, T. Gudermann, *Am. J. Physiol. Renal Physiol.* **286**, F599 (2004).
17. Mutations in PPARγ and Akt2 may be an exception, as the few patients reported have both insulin resistance and hypertension.
18. Materials and methods are available as supporting material on *Science* Online.
19. MITOMAP: A Human Mitochondrial Genome Database, available at www.mitomap.org.
20. M. Sprinzl, C. Horn, M. Brown, A. Ioudovitch, S. Steinberg, *Nucleic Acids Res.* **26**, 148 (1998).
21. S. H. Kim *et al.*, *Science* **179**, 285 (1973).
22. G. J. Quigley, A. Rich, *Science* **194**, 796 (1976).
23. S. S. Ashraf *et al.*, *RNA* **5**, 188 (1999).
24. D. C. Wallace, *Science* **283**, 1482 (1999).
25. P. Maechler, C. B. Wollheim, *Nature* **414**, 807 (2001).
26. K. F. Petersen *et al.*, *Science* **300**, 1140 (2003).
27. V. K. Mootha *et al.*, *Nature Genet.* **34**, 267 (2003).
28. A. Trifunovic *et al.*, *Nature* **429**, 417 (2004).
29. V. Carelli, C. Giordano, G. d'Amati, *Trends Genet.* **19**, 257 (2003).
30. R. F. Reilly, D. H. Ellison, *Physiol. Rev.* **80**, 277 (2000).
31. D. B. Simon *et al.*, *Nature Genet.* **12**, 24 (1996).
32. R. A. Zager, A. C. Johnson, S. Y. Hanson, *Am. J. Physiol. Renal Physiol.* **285**, F1092 (2003).
33. A. Atlante *et al.*, *Int. J. Mol. Med.* **1**, 709 (1998).
34. R. S. Vasan *et al.*, *JAMA* **287**, 1003 (2002).

35. We thank the members of K129 for their generous participation in this project; I. Beerman, C. Mendenhall, and F. Niazi for assistance with patient evaluation; D. Befroy and S. Dufour for assistance with spectroscopy; C. Ariyan and J. Kim for help with muscle biopsy; C. Garganta for measurement of urinary amino acids and organic acids; the staff of the Yale General Clinical Research Center; and A. Gharavi for helpful discussions. Supported by NIH grant nos. MO1 RR-00125, P50 HL-55007, and R01 DK-49230. A.H. is the recipient of an American Heart Association Fellowship (no. 0475003N).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1102521/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
Tables S1 to S3
References and Notes

8 July 2004; accepted 6 September 2004
Published online 21 October 2004;
10.1126/science.1102521
Include this information when citing this paper.

Multidimensional Drug Profiling By Automated Microscopy

Zachary E. Perlman,[1,2*] Michael D. Slack,[3*†] Yan Feng,[1*‡] Timothy J. Mitchison,[1,2] Lani F. Wu,[3§] Steven J. Altschuler[3§]

We present a method for high-throughput cytological profiling by microscopy. Our system provides quantitative multidimensional measures of individual cell states over wide ranges of perturbations. We profile dose-dependent phenotypic effects of drugs in human cell culture with a titration-invariant similarity score (TISS). This method successfully categorized blinded drugs and suggested targets for drugs of uncertain mechanism. Multivariate single-cell analysis is a starting point for identifying relationships among drug effects at a systems level and a step toward phenotypic profiling at the single-cell level. Our methods will be useful for discovering the mechanism and predicting the toxicity of new drugs.

High-throughput methods for describing cell phenotype such as transcriptional and proteomic profiling allow broad, quantitative, and machine-readable measures of the responses of cell populations to perturbation (*1*–*4*). Automated microscopy has the potential to complement these profiling approaches by allowing fast and cheap collection of data describing protein behaviors and biological pathways within individual cells (*5*–*9*). Accessing these data to produce useful profiles of cell phenotype will require new image

analysis methods, the development of which has so far lagged behind the adoption of high-throughput imaging technologies.

In the context of drug discovery, profiling technologies are useful in measuring both drug action on a desired target in the cellular milieu and drug action on other targets. Ideally, such profiling should be performed as a function of drug concentration, because several factors make the effects of drugs highly dose dependent. For example, the degree to which a primary target is perturbed may affect different downstream pathways differently, and drugs can bind to multiple targets with different affinities. In some cases, the therapeutic mechanism may involve binding to more than one target with differing affinity (*10*, *11*). To date, drug effects have been broadly profiled with transcript analysis, proteomics, and measurement of cell line dependence of toxicity (*11*–*21*). In these studies, multidimensional profiling methods were only applied at a single-drug concentration. The only studies in which drug dose has been explicitly considered as a variable used the degree of cell proliferation, an essentially one-dimensional (1D) readout of phenotype (*12*, *13*). Two recent reviews have highlighted the possibility of using combinations of targeted phenotypic imaging screens to generate profiles of drug activity (*6*, *22*). Here, we suggest that large sets of unbiased measurements might serve as high-dimensional cytological profiles analogous to transcriptional profiles. We present a method based on hypothesis-free molecular cytology that provides multidimensional single-cell phenotypic information yet is simple and inexpensive enough to allow extensive dose-response profiles for many drugs.

We assembled a test set of 100 compounds (table S1). Of these, 90 were drugs of known mechanism of action, six were blinded alternate titrations from this set of known drugs, one (didemnin B) was a toxin reported to have multiple biological targets (*23*), and three were drugs of unknown mechanism. The known drug set was chosen to cover common mechanisms of toxicity or therapeutic action in cancer and other diseases and to include several groups with a common target (macromolecule or pathway) but unrelated structures. We analyzed 13 threefold dilutions of each drug, covering a final concentration range on cells from micromolar to picomolar [table S2 and supporting online material (SOM) text A]. HeLa (human cancer) cells were cultured in 384-well plates to near confluence, treated with drugs for 20 hours, fixed, and stained with fluorescent probes for various cell components and processes. We chose 11 distinct probes that covered a range of cell biology, multiplexing a DNA stain and two antibodies per well [the probe sets are SC35, anillin; α-tubulin, actin; phospho-p38, phospho–extracellular signal–regulated kinase (ERK); p53, c-Fos; phospho–adenosine 3′,5′-monophosphate response element–binding protein (CREB), calmodulin]. Using automated fluorescence microscopy, we collected images of up to ~8000 cells from each well. On each plate, 26 wells were treated only with dimethyl sulfoxide (DMSO) to generate a control population (SOM text A). The experiment was performed twice in parallel to provide a replicate data set. Image segmentation procedures were used to automatically identify nuclei and nuclear organelles, and cytoplasmic regions were approximated as an annulus surrounding each identified nucleus (Fig. 1A and SOM text B). For each cell, region, and probe, a set of descriptors was measured. These included measures of size, shape, and intensity, as well as ratios of intensities between regions (93 descriptors total, table S3). In all, $\sim 7 \times 10^7$ individual cells were identified from >600,000 images, yielding $\sim 10^9$ data points.

We can examine the population response of each descriptor to increasing concentrations of a given drug, which we show with the genotoxic compound camptothecin (*24*) (Fig. 1B). At low concentrations, the histogram for the total DNA content has the characteristic bimodal shape reflecting a mixture of G_1, S, and G_2/M cell populations. G_2 and M populations may be distinguished by 2D display of total DNA signal against nuclear area (*25*). As drug concentration increases, the cells arrest with S/G_2 DNA content (*24*). The measured DNA content distribution shifts leftward as dose increases, and at the highest concentrations apoptosis is widely induced. Anillin, a cytokinesis protein whose levels reflect cell cycle progression

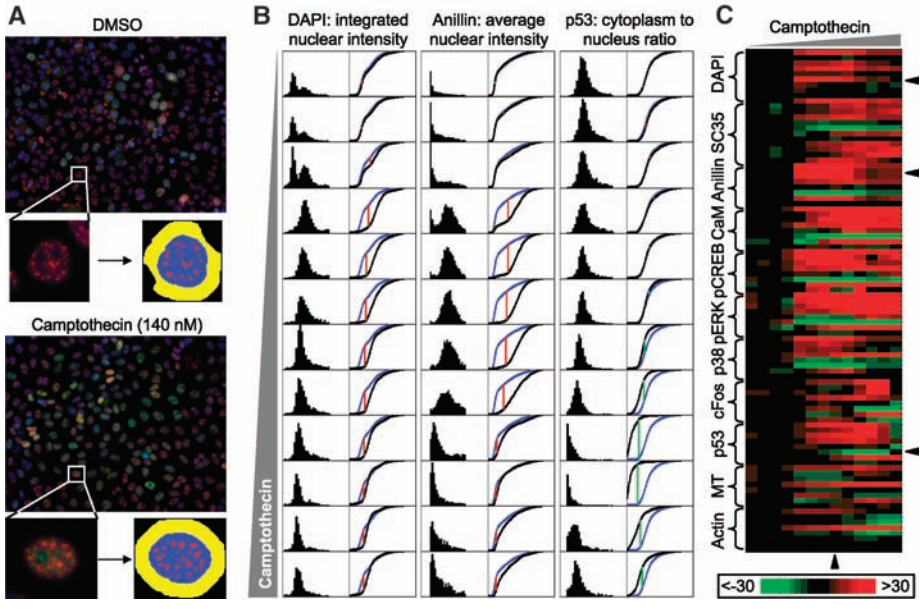


Fig. 1. Key steps in algorithm for reducing image data to compound profile. (**A**) Image segmentation. For each image [examples show DNA (blue), SC35 (red), and anillin (green)], we generated a nuclear region (blue) and a set of associated regions [shown here are cytoplasmic annulus (yellow) and SC35 speckles (red)]. For each defined nuclear region, we measure multiple descriptors. (**B**) Quantification of population response. For a given compound, titration, and descriptor, we generated a population histogram and related cumulative distribution function (cdf) (black) to be compared with the control population (blue). Shown is a threefold dilution series ranging from 65 pM to 35 μM camptothecin. We reduced each experimental cdf to a single dependent variable through comparison with a control population with the nonparametric KS statistic against a control population (SOM text C). Each vertical red or green line indicates the position and sign of the maximal height difference between the curves; this height is the KS statistic. (**C**) Heat map of compound profile. A *z* score is calculated for each KS statistic (SOM text C), and the vector of *z* scores for all descriptors and all titrations is displayed for rapid visual assessment. Increased scores are represented in red and decreased in green, with intensity encoding magnitude. Arrowheads to the right indicate descriptors shown in (B), and the arrowhead at the bottom indicates the dose shown in (A).

[1]Institute of Chemistry and Cell Biology, Harvard Medical School, Boston, MA 02115, USA. [2]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [3]Bauer Center for Genomics Research, Harvard University, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†Present address: Alphatech, Inc., San Diego, CA 92123, USA.
‡Present address: Novartis Institutes for BioMedical Research, Cambridge, MA 02139, USA.
§To whom correspondence should be addressed. E-mail: saltschuler@cgr.harvard.edu (S.J.A.); lwu@cgr.harvard.edu (L.F.W.)

(*26*), shows marked nuclear accumulation in the G_2 arrested state (Fig. 1A). p53, a transcription factor that is part of the genotoxic response pathway, is strongly induced at high camptothecin concentrations, but much less so at concentrations sufficient to promote G_2 arrest (Fig. 1B).

For profiling studies, it is useful to reduce each population of descriptor values to a single number. Our study made several demands of this reduction: It must be able to compare distributions of arbitrary shape (Fig. 1B); it must be robust to variation in dynamic range and noise levels among different descriptors; it must convert different types of measurement into a common unit for comparison; it must be descriptor parameterization independent (e.g., an intensity ratio should behave the same as its reciprocal); and it must be insensitive to the precise quantitative relationship between antibody-staining intensity and antigen density. We devised a measure based on the Kolmogorov-Smirnov (KS) statistic, allowing nonparametric comparison of experimental and control distributions from the same plate (Fig. 1B, fig. S1, and SOM text C). Dividing by a measure of the variability within the control population yielded a *z* score, which can be displayed as a function of descriptor and drug concentration in a heat plot to allow rapid visual comparison of compound response profiles (Fig. 1C). These plots represent a family of dose-response curves for a single drug but differ from traditional curves reflecting changes in a biochemical measurement. In particular, the relationship between *z* score and the original physical measure may be nonlinear. For example, the statistically significant responses of p53 to low doses of camptothecin (Fig. 1C) reflect subtle effects not easily discerned by eye in the source images.

The heat plots typically have a sharp transition, reflecting a concentration at which many descriptors become different from control values. We will refer to this as the primary effective concentration (PEC) for the drug. The isolated responses observed at some low concentrations represent noise that could be reduced by increasing replicates, improving experimental procedures, and normalizing for local variation in cell density. For 39 drugs, we saw no strong effect, leaving a heat plot dominated by noise. Those drugs either lack a target in HeLa cells, were used at inactive dosages, or effected changes not detectable with our antibody set. For almost all of the 61 drugs that showed a strong response, some descriptors responded at concentrations other than the PEC (Fig. 2). This may reflect varying biological consequences of low and high saturation of a single target, or it may reflect interactions with multiple targets with different affinities. For example, camptothecin binds primarily to DNA complexes with topoisomerase I, promoting DNA strand breaks and S-phase arrest at low concentrations, but it also blocks transcription and a number of other cellular processes at higher concentrations (*24*). Other drugs in our test set are known to have multiple targets, such as histone deacetylase inhibitors (*27*) and the general kinase inhibitor staurosporine (*28*) and were thus expected to show complex dose-response behavior. Such phenotypic

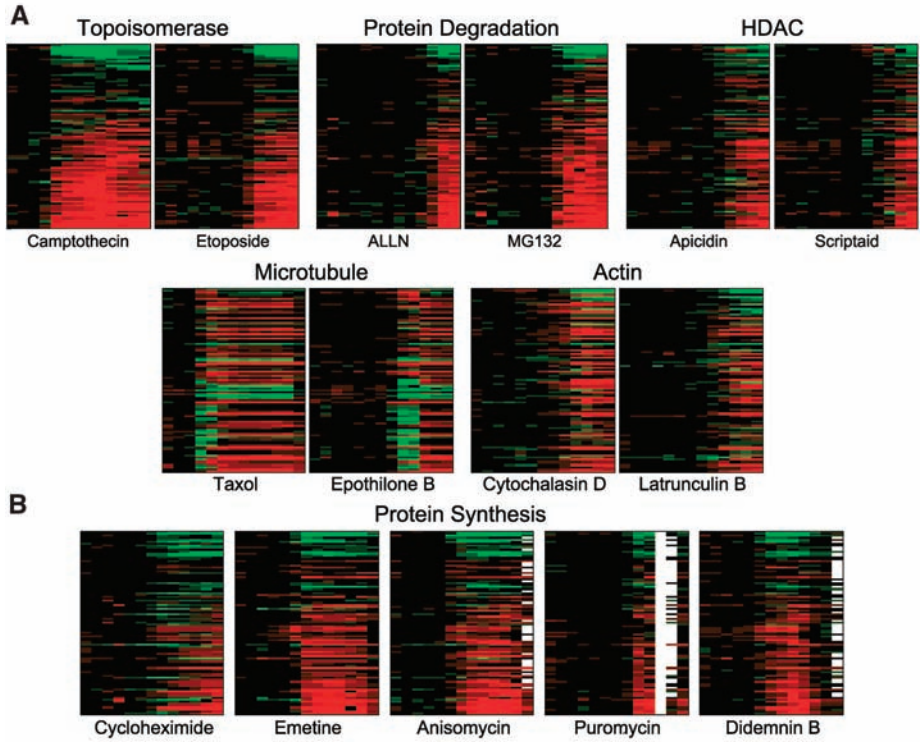


Fig. 2. Comparison of compound profiles. As in Fig. 1C, the *x* axis shows increasing dose and the *y* axis encodes descriptors. Dose ranges are shown from 65 pM to 35 μM for all drugs except epothilone B, which is shown from .65 pM to .35 μM. Color scale is as in Fig. 1C. For ease of visualization, descriptors in all profiles are sorted in decreasing order of camptothecin response. (**A**) Compounds of similar mechanism show similar profiles. Shown are representative compound profiles. HDAC, histone deacetylase; ALLN, *N*-acetyl-Leu-Leu-norleucinal. (**B**) Compound profiles can distinguish differences between drugs with similar mechanisms. Wells with too few cells for analysis are represented in white.

Table 1. Assessment of TISS by literature categories. For each category that has more than two compounds, we computed two sets of TISS scores: pairwise TISS comparisons between members of the category (intrapair) and comparisons in which only one element of the pair is in the category (interpair). As a crude in silico comparison to other cell-based assays such as fluorescence-activated cell sorting (single-cell based) and cytoblots (whole-population based), we repeated this procedure with a descriptor set consisting of only total intensity measures and compared it with either our KS-based TISS values or a mean-based TISS values (SOM text C). *P* values (columns 2 to 4) describe the probability that the rank ordering of the two sets of TISS values would have been seen by random draws from the same distribution (SOM text C). KS, KS-based TISS (*P* value); mean, mean-based TISS (*P* value).

Category	All descriptors	Total intensity descriptors		No. pairwise TISS comparisons	
	KS	KS	Mean	Intrapair	Interpair
Actin	0.025	0.776	0.327	6	218
DNA replication	0.011	0.057	0.007	3	168
Histone deacetylase	0.001	0.024	0.489	10	265
Kinase	0.223	0.746	0.902	3	168
Kinase CDK	0.057	0.221	0.050	6	218
Microtubule	3.86×10^{-20}	9.81×10^{-6}	0.295	55	484
Protein synthesis	6.02×10^{-5}	0.004	0.180	15	309
Topoisomerase	0.005	0.011	0.693	3	168
Vesicle trafficking	0.206	0.314	0.514	3	168

complexity may help explain why toxicity at high doses is common even for therapeutic drugs that are apparently highly selective at the level of target binding.

Drugs with common reported targets but diverse chemical structures often showed similar profiles readily distinguished from those of drugs of different mechanism (Fig. 2A). In other cases, markedly different profiles were evident within a family, most notably the protein synthesis inhibitors (Fig. 2B). This may reflect different cell responses to alternative biochemical mechanisms of poisoning ribosomes (*29*) or perhaps the existence of alternate targets (*23*).

When comparing drug mechanism, changes in specificity (and thus phenotype) are relevant, but changes in affinity (and thus PEC) are not. Two different dosage series of the same drug should result in similar heat plots shifted along the concentration axis. We developed a titration-invariant similarity score (TISS) to allow comparison between dose-response profiles independent of starting dose (SOM text C). TISS values were generated for the 61 compounds that showed significant signal, and these were used for unsupervised clustering (Fig. 3). TISS was successful at grouping compounds with similar reported targets (Table 1). As expected, clustering reflected biological mechanism rather than chemical similarity. For example, kinase inhibitors, most of which are adenosine 5′-triphosphate–mimetic compounds, did not cluster as a group. Clustering was poor even within a set of kinase inhibitors with overlapping targets [cyclin-dependent kinase (CDK) inhibitors], perhaps reflecting variable inhibition of other kinases. The CDK inhibitors related by structure and reported target (purvalanol, roscovitine, and olomoucine) did cluster.

Of the blinded alternate titrations of known drugs, scriptaid, hydroxyurea, emetine, and two alternate series of nocodazole showed significant responses. These clustered closely with their unblinded counterparts and compounds of similar reported mechanism. Didemnin B, for which the reported range of activities includes inhibition of protein synthesis (*23*), clustered with ribosome inhibitors (Fig. 2B). Two of the three poorly characterized compounds showed strong responses. One, concentramide, is difficult to interpret. The other, austocystin, clusters with transcription and translation inhibitors. Preliminary experiments suggest that this compound inhibits transcription in vitro (*25*). Thus, our methods can group compounds of like mechanism and thereby suggest mechanism for new drugs.

Extensions of cytological profiling to reflect dependencies among descriptors will allow more sophisticated analysis of drug

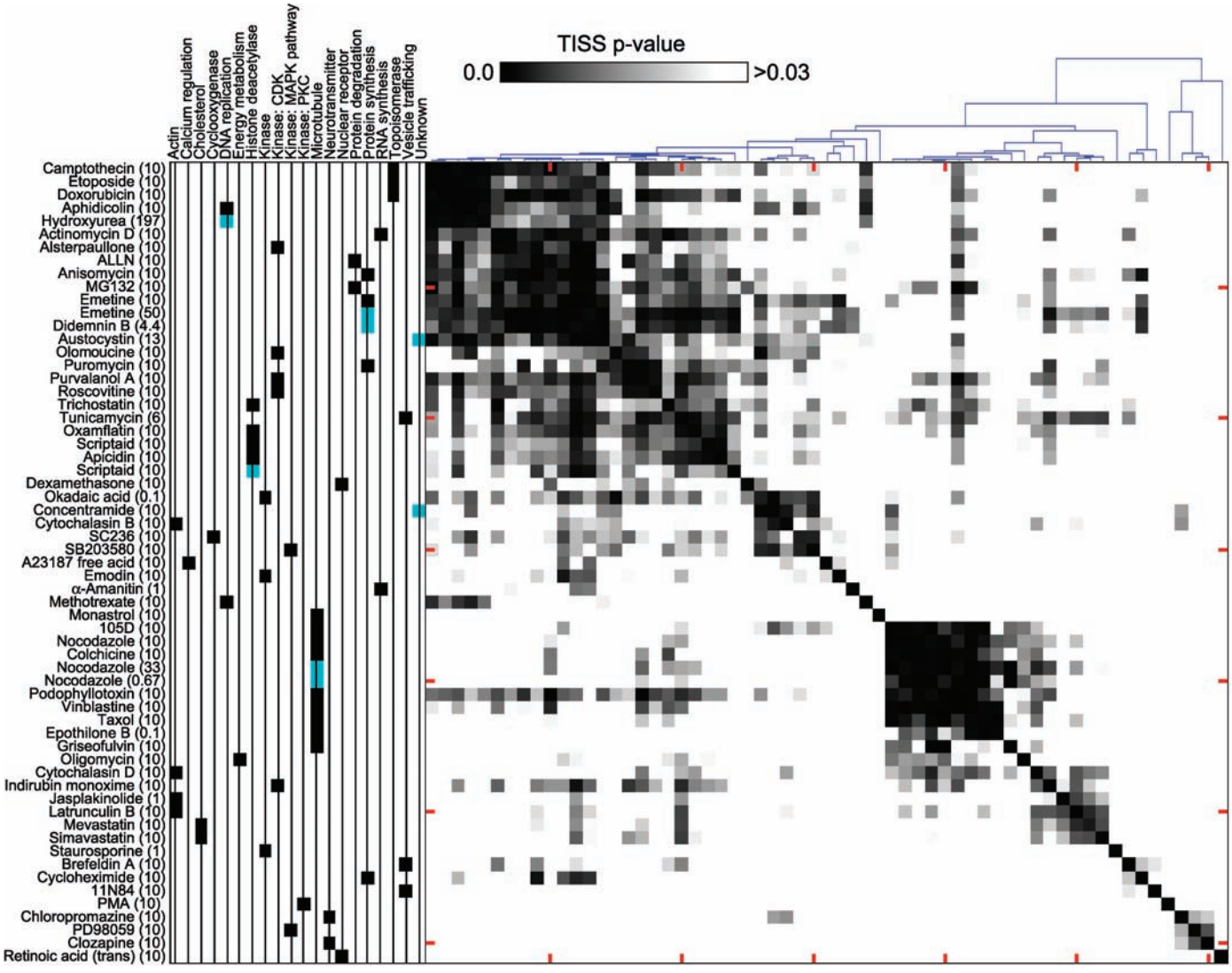


Fig. 3. Hierarchical clustering of the 61 most responsive compound profiles by TISS values. Compound stock concentrations (μM) are in parentheses (fig. S3). Left panel shows mechanism of compound as described in literature. In blue are compounds that were blinded or are of unknown mechanism. Middle panel shows matrix of *P* values derived from pairwise TISS values (SOM text C). Dendrogram at top shows degree of association.

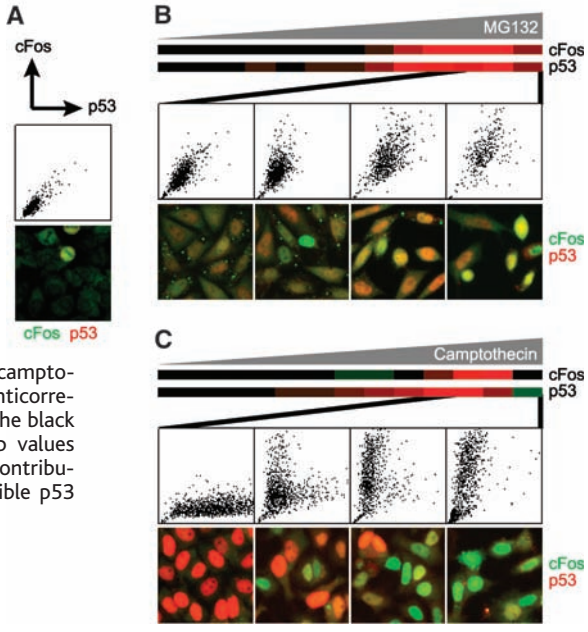


Fig. 4. Single-cell analysis shows differing patterns of dose-dependent p53 and c-Fos responses to different drugs. (**A**) Scatter plot of average nuclear p53 intensity versus average c-Fos intensity in a typical control well and representative image. The bright cells at the top of the image are in mitosis. (**B**) Dose-dependent increases in response to MG132 shown in heat maps are correlated in scatter plots and images (orange nuclei), shown for the four highest concentrations. (**C**) Dose-dependent increases in response to camptothecin shown in heat maps are anticorrelated in scatter plots and images. The black (c-Fos) and green (p53) heat map values for the highest dose reflect the contribution of apoptotic cells with negligible p53 and c-Fos nuclear staining.

responses at a systems level. For example, both p53 and c-Fos, a transcription factor involved in mitogen-activated protein kinase (MAPK) signaling, are involved in cell stress responses, but the interrelationship of the p53 and MAPK pathways is poorly understood (*30*). Single-cell profiling reveals that different drug mechanisms induce different relative patterns of response by these two pathways (Fig. 4). The proteasome inhibitor MG132 causes increased correlated induction in these pathways, whereas responses to camptothecin are anticorrelated. Anticorrelated responses observed in fixed-time images may reflect switching of mutually exclusive cell states in response to different degrees of stress or might reflect a dynamic temporal response, such as oscillation, that is not synchronized among cells (*31*). These data help establish a concentration and time window, but live imaging will be required to distinguish between these hypotheses.

Cytometric dose-response profiling is a fast and cheap method for quantitatively surveying broad ranges of individual cell responses. We have used our methods to assign mechanism to blinded and uncharacterized drugs and to suggest systems-level relationships between signaling pathways. The complex dose-response curves and large cell-to-cell variability we frequently observed reinforce the utility of unbiased multidimensional characterization of drug effects over wide ranges of doses.

Many improvements and extensions of this work are possible. These include better lab automation, broader drug reference sets, different types of perturbation (such as RNA interference), improved strategies for cell segmentation, more sophisticated feature extraction (*5*, *9*), different sets of antibody probes and cells, the inclusion of more time points and live cell imaging, and the integration of complementary profiling strategies. Additionally, our methods may be extended to allow the characterization of responses by subpopulations defined by such variables as cell cycle state, cell density, or neighboring environment. This analysis, extended to work in tissues or clinical samples, offers the potential to speed the identification of toxic compounds during therapeutic drug development and the targeting of drug effects to specific subtypes of cells.

References and Notes

1. M. B. Eisen, P. T. Spellman, P. O. Brown, D. Botstein, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14863 (1998).
2. A. C. Gavin *et al.*, *Nature* **415**, 141 (2002).
3. Y. Ho *et al.*, *Nature* **415**, 180 (2002).
4. P. Uetz *et al.*, *Nature* **403**, 623 (2000).
5. R. F. Murphy, M. Velliste, G. Porreca, *J. VLSI Signal Process.* **35**, 311 (2003).
6. J. H. Price *et al.*, *J. Cell. Biochem. Suppl.* **39**, 194 (2002).
7. W. K. Huh *et al.*, *Nature* **425**, 686 (2003).
8. T. L. Saito *et al.*, *Nucleic Acids Res.* **32**, D319 (2004).
9. C. Conrad *et al.*, *Genome Res.* **14**, 1130 (2004).
10. J. G. Hardman, L. E. Limbird, A. G. Gilman, Eds., *The Pharmacological Basis of Therapeutics* (McGraw-Hill, New York, ed. 10, 2001), pp. 39–42.
11. M. J. Marton *et al.*, *Nature Med.* **4**, 1293 (1998).
12. J. N. Weinstein *et al.*, *Science* **275**, 343 (1997).
13. K. D. Paull, C. M. Lin, L. Malspeis, E. Hamel, *Cancer Res.* **52**, 3892 (1992).
14. U. Scherf *et al.*, *Nature Genet.* **24**, 236 (2000).
15. E. C. Gunther, D. J. Stone, R. W. Gerwien, P. Bento, M. P. Heyes, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9608 (2003).
16. D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nature Biotechnol.* **21**, 687 (2003).
17. M. A. Lindsay, *Nature Rev. Drug Discov.* **2**, 831 (2003).
18. P. Y. Lum *et al.*, *Cell* **116**, 121 (2004).
19. G. Giaever *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 793 (2004).
20. S. J. Haggarty, P. A. Clemons, S. L. Schreiber, *J. Am. Chem. Soc.* **125**, 10543 (2003).
21. D. E. Root, S. P. Flaherty, B. P. Kelley, B. R. Stockwell, *Chem. Biol.* **10**, 881 (2003).
22. V. C. Abraham, D. L. Taylor, J. R. Haskins, *Trends Biotechnol.* **22**, 15 (2004).
23. M. D. Vera, M. M. Joullie, *Med. Res. Rev.* **22**, 102 (2002).
24. C. J. Thomas, N. J. Rahier, S. M. Hecht, *Bioorg. Med. Chem.* **12**, 1585 (2004).
25. Z. E. Perlman *et al.*, data not shown.
26. C. M. Field, B. M. Alberts, *J. Cell Biol.* **131**, 165 (1995).
27. M. Yoshida *et al.*, *Cancer Chemother. Pharmacol.* **48** (suppl. 1), S20 (2001).
28. M. E. Noble, J. A. Endicott, L. N. Johnson, *Science* **303**, 1800 (2004).
29. J. D. Laskin, D. E. Heck, D. L. Laskin, *Toxicol. Sci.* **69**, 289 (2002).
30. B. Kaina, *Biochem. Pharmacol.* **66**, 1547 (2003).
31. G. Lahav *et al.*, *Nature Genet.* **36**, 147 (2004).

32. We thank A. Daneau, M. Ethier, and B. Mantenuto at the Bauer Center for Genomics Research for their assistance with the use of the Bauer Center computational cluster and K. Maciag, A. Murray, O. Rando, A. Salic, and A. Yonetani for helpful discussions. Supported in part by National Cancer Institute PO1 CA078048. Z.E.P. is a Howard Hughes Medical Institute Predoctoral Fellow.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5699/1194/DC1
SOM Text
Figs. S1 to S5
Tables S1 to S4
Database S1

25 May 2004; accepted 8 September 2004

New Products

Millipore
For more information
800-MILLIPORE
www.millipore.com
http://science.labvelocity.com

Kinase Assay Plates

MultiScreenHTS-PH filter plates for high-throughput kinase screening assays are automation-compatible 96-well plates that follow a prevalidated protocol for high sensitivity, high specificity, and reliability. For optimized kinase assay performance, the plates incorporate PH paper, which retains peptide substrates to remove unreacted, unlabeled adenosine triphosphate. The PH paper also provides lower background and improved discrimination between drug target hits and non-hits. In addition, the paper is not subject to naturally fluorescent or quenching compounds. For this reason, the plates are especially compatible with aromatic-based libraries. Kinase assays are a critical step in both lead identification and characterization. Compared with most homogeneous assays kinase assays, completed on MultiScreenHTS-PH filter plates are easy to set up.

Applied Kilovolts
For more information
410-308-0069
www.appliedkilovoltsms.com
http://science.labvelocity.com

High-Voltage Power Supplies

These precision high-voltage power supplies offer performance enhancement options for mass spectrometer producers. Developed to minimize both direct and injected ripple, the HP Series power supplies can ensure maximum sensitivity and resolution while maintaining extreme-temperature stability and long-term reliability. They are available in a variety of configurations and power supply ranges to meet specific engineering requirements. From 500V to 30kV, the HP, HPR, HPF, and HPRD Series for mass spectrometers and floating detector applications offer precision unipolar and bipolar high-voltage supplies with low output ripple.

Niles Scientific
For more information
415-346-6461
www.spotreader.com
http://science.labvelocity.com

Microarray Image Processing Software

SpotReader 1.0 is new microarray image processing software for spotted (robot printed) images. Major features include hands-off creation of data analysis grids, a mini-database to organize data, and a significant advance in a grid editing user interface. Automation features allow walk-away processing of microarray images: if you have your scanner save images into a folder, grids are made automatically. The true Windows interface provides live-update automatic flagging, a pop-up magnifying glass, modeless editing, and a feature-pinning function for extremely hard-to-grid arrays.

CEM Corp
For more information
704-821-7015
www.cem.com
http://science.labvelocity.com

Microwave System for Lower Temperatures

The Discover CoolMate microwave system is designed to perform sub-ambient temperature chemistries. Discover CoolMate performs temperature-sensitive reactions including carbohydrate chemistry and the formation of carbanions and other highly reactive intermediates easily and controllably. The system also allows the safe use of highly reactive species, which would normally be excluded when using a microwave system. The system's proprietary, jacketed, low-temperature vessel and cooling technology keep the bulk temperature of the reactants low, preventing thermal degradation of the product while microwave energy drives the reaction kinetically. Temperature measurement is accomplished with an in situ fiber-optic probe to insure fast, accurate, and convenient reaction feedback control.

Greiner Bio-One
For more information
800-884-4703
www.gbo.com
http://science.labvelocity.com

Low Birefringence Microplates

As a key technology in structural genomics and structure-based drug development, protein crystallization requires increasingly efficient analytical procedures. Polarized light is now used with imaging systems to better distinguish protein crystal structures from precipitate artifacts. Polystyrene, the material widely used for microplate manufacture, exhibits a wide spectrum of color to interfere with polarized detection. To reduce this kind of background interference, the CrystalQuick-LBR Plates for sitting-drop vapor diffusions are manufactured of a low birefringent resin with superior optical properties for crystal scoring. Characterized by exceptional transparency, high chemical resistance, and low water absorption, the plates feature three crystallization wells per each 96-well reservoir. Available with "square" flat bottom or "round" curved bottom crystallization wells, the plates facilitate high-throughput screening in manual or fully automated systems.

Textco BioSoftware
For more information
603-643-1471
www.textco.com
http://science.labvelocity.com

Gene Inspector Software

Gene Inspector 1.6 for Macintosh OS X integrates a comprehensive sequence analysis package and powerful illustration tools into a versatile electronic research notebook. The intuitive interface allows users to organize large amounts of diverse yet interrelated information, enabling researchers to plan, annotate, and execute experiments and analyses in a single application. The notebook features help users easily produce documents to summarize and manage complex experiments and produce publication-quality illustrations. The analysis setups allow parameters to be set and stored so consistent results are easy to obtain, even among multiple users. The program features more than 60 standard analyses for both nucleic acids and peptides, including multiple sequence alignments, BLAST searching, dot matrix comparisons, and prediction of the physical properties of translation-product peptides.

Axxora
For more information
858-550-8830
www.axxora.com
http://science.labvelocity.com

Literature

Fluorescent Proteins is a 16-page guide describing the use of novel fluorescent proteins in various biological activities. Among the proteins featured are KFP-Red, a photoactivatable Kindling-Red fluorescent protein; CopGFP, a bright and fast-maturing green fluorescent protein; PhyYFP, a true yellow fluorescent protein; and t-HcRed, a non-oligomeric far-red fluorescent protein. This reference details a variety of in vivo applications, including gene expression analysis, protein localization, cell labeling, and protein–protein interaction and co-localization studies. It discusses the general properties, advantages, limitations, and recommended uses of these innovative fluorescent proteins.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and government organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier by visiting http://science.labvelocity.com on the Web, where you can request that the information be sent to you by e-mail, fax, mail, or telephone.

Published by AAAS